(8)

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

SELECTING A SUBSET OF STIMULUS-RESPONSE
PAIRS WITH
MAXIMAL TRANSMITTED INFORMATION

by

Michael J. Sheehan

March, 1992

Principal Advisor: Eric S. Theise

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL *(If applicable)* 55	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS *(City, State, and ZIP Code)* Monterey, CA 93943-5000	7b. ADDRESS *(City, State, and ZIP Code)* Monterey, CA 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, and ZIP Code)*	10. SOURCE OF FUNDING NUMBERS Program Element No. / Project No. / Task No. / Work Unit Accession Number

11 TITLE *(Include Security Classification)*
SELECTING A SUBSET OF STIMULUS-RESPONSE PAIRS WITH MAXIMAL TRANSMITTED INFORMATION

12 PERSONAL AUTHOR(S) Sheehan, Michael, James

13a. TYPE OF REPORT	13b. TIME COVERED	14 DATE OF REPORT *(year, month, day)*	15. PAGE COUNT
Master's Thesis	From To	March 1992	136

16 SUPPLEMENTARY NOTATION
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 COSATI CODES			18. SUBJECT TERMS *(continue on reverse if necessary and identify by block number)*
FIELD	GROUP	SUBGROUP	information, information thoery, transmitted information, confusion, recognition, discriminability, stimulus-response, human performance, man-machine, mathematical programming

19 ABSTRACT *(continue on reverse if necessary and identify by block number)*

System designers are often faced with the task of assigning symbolic representations to user actions, e.g., icons to choices in graphical interfaces. When a confusion matrix--on discriminability of symbols--is available, it is used to guide the selection of the set of symbols to be implemented. While trial and error methods or clustering approaches have been used to analyze this problem, it was only recently that a true optimization approach was offered. Theise (1989) formulated the symbol selection problem as a zero-one integer programming problem whose objective function was linked to the minimization of within-subset confusion.

Confusion is not the traditional metric used by human factors engineers to analyze confusion matrices. Rather, transmitted-information--a metric from information theory--has long been used to evaluate system performance. The purpose of this thesis is to formulate a model of subset selection in which transmitted information will be maximized.

It is possible to specify a correct model, although current algorithms are incapable of solving it. This thesis reports on the performance of a GAMS-based approximation to the original model, as well as an exhaustive enumeration scheme. Solutions from both information-theoretic approaches are compared to solutions from the confusion/recognition model.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS REPORT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Eric S. Theise	22b. TELEPHONE *(Include Area code)* 646-3215 / 22c. OFFICE SYMBOL ORTH

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

SELECTING A SUBSET OF STIMULUS-RESPONSE PAIRS WITH MAXIMAL TRANSMITTED INFORMATION

by

Michael J. Sheehan
Captain, United States Air Force
B.S., Wright State University, 1986

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS TECHNOLOGY

from the

NAVAL POSTGRADUATE SCHOOL
March 1992

Author: ______________________
Michael J. Sheehan

Approved by: ______________________
Eric S. Theise, Principal Advisor

____
Gary K. Poock, Associate Advisor

____
Carl R. Jones, Chairman
Command, Control, and Communications Academic Group

ABSTRACT

System designers are often faced with the task of assigning symbolic representations to user actions, e.g., icons to choices in graphical interfaces. When a *confusion matrix*--on discriminability of the symbols--is available, it is used to guide the selection of the set of symbols to be implemented. While trial and error methods or clustering approaches have been used to analyze this problem, it was only recently that a true optimization approach was offered. Theise (1989) formulated the symbol selection problem as a zero-one integer programming problem whose objective function was linked to the minimization of within-subset confusion.

Confusion is not the traditional metric used by human factors engineers to analyze confusion matrices. Rather, *transmitted-information*--a metric from information theory--has long been used to evaluate system performance. The purpose of this thesis is to formulate a model of subset selection in which transmitted information will be maximized.

It is possible to specify a correct model, although current algorithms are incapable of solving it. This thesis reports on the performance of a GAMS-based approximation to the original model, as well as an exhaustive enumeration scheme. Solutions from both information-theoretic approaches are compared to solutions from the confusion/recognition model.

TABLE OF CONTENTS

I. INTRODUCTION

A. PURPOSE FOR THESIS

The problem presented in this thesis was introduced to the author by Dr. Eric S. Theise as a follow-up to a paper he had published in *Human Factors* in 1989 titled "Finding a Subset of Stimulus-Response Pairs with Minimum Total Confusion: A Binary Integer Programming Approach." As the title implies, the paper dealt with optimization models using binary integer programming. The idea was to select an optimal subset from a given set of stimulus-response (S-R) pairs using confusion as a guiding index to optimality. Dr. Theise was interested in further research into optimal subsets; however, he was interested in using information theory to develop a guiding index rather than using confusion.

A brief introduction to S-R pairs and their use in confusion matrices is warranted here. An S-R pair is simply a stimulus and the corresponding response to that stimulus. A confusion matrix can be formed from stimulus-response experimentation. An example of a confusion matrix taken from Clarke's (1957) work on phonetic syllables is presented in Table 1. The matrix is formed by presenting a test subject with a stimulus such as the syllable *ka*. If the test subject correctly identifies the syllable as *ka*, a tally is made on given subspecialty are typically not aware of optimizing techniques being used in other subspecialties that could be of potential benefit to them. (Fisher, in press) The research in this paper is aimed at using operations research methods to solve a problem of an optimal performance nature from the realm of human factors. As such, the purpose of this paper is to produce an optimization model that will select a subset of S-R pairs from a given set S-R pairs with the objective of maximizing *transmitted-information*. Appropriately, this model will be referred to as the *Transmitted-Information Model*.

In a military environment, this research has implications for the command, control, and communications (C3) discipline. C3 can often be the deciding factor in the failure or success of military missions. This type of research can help system designers make C3 systems more user-friendly through better human-system interfaces, thus helping the commander achieve his goals more effectively. Other areas that may benefit from this type of research include antisubmarine warfare (ASW), computer science including software design, and human-system interface applications such as aircraft cockpit design.

B. RESEARCH QUESTIONS

The answers to several questions are explored in this paper. The questions of interest are as follows: Can a model be formulated that uses an information theoretic framework to select a subset of S-R pairs in such a way as to maximize the amount of information transmitted? Can this model be solved using standard mathematical programming software? If not, can a special purpose algorithm or effective heuristic be developed? How does the solution to this model compare with the minimal confusion solution for the same confusion matrix data?

C. SCOPE AND ORGANIZATION

What this paper attempts to do is lay the groundwork for better empirical optimization in problems dealing with human factors. This can be extremely beneficial to the C3 community when working on problems involving the human-system interface, especially when time is critical, and mistakes can cost lives and possibly jeopardize national security.

In the process of laying this groundwork, a model will be developed that will optimize the transmitted-information from a subset of S-R pairs. The results of the application of the model to 17 data sets will be compared to the results from the model previously developed by Theise (1989). The comparison will attempt to determine the better optimization method.

This thesis is broken into seven chapters. Chapter I provides the purpose, scope, and organization of the thesis. Chapter II explores some background in the human-system interface area with special attention to C3 issues.

Chapter III will provide background on the previous work by Theise (1989) and will define some of the concepts to be used throughout the thesis. Chapter IV introduces information theory and its associated terms and concepts to be used in developing a new optimization model. Chapter V presents the concept of optimal subsets using information theory. In this chapter, the optimization model is developed, and is then applied to 17 available data sets.

Chapter VI provides an analysis of the results produced in Chapter V and compares these results to the results of the same data applied to the confusion/recognition model. Finally, Chapter VII presents conclusions and recommendations including areas that may warrant further study.

II. BACKGROUND

A. NATURE OF THE PROBLEM

1. Human Factors Defined

The field of human factors is concerned with improving the interface between people and machines or objects. For this reason, human factors is often referred to by the more descriptive term--human-system interface.

> Human factors, then, seeks to change the things people use and the environments in which they use these things to better match the capabilities, limitations, and needs of people. (Sanders and McCormick, 1987, p. 4)

With this in mind, it should be obvious that a primary goal of human factors is to improve the efficiency and effectiveness of people in the performance of the various tasks required of them.

2. Optimal System Design

System designers are not always trained in human factors engineering and, therefore, do not think in terms of optimal performance. Instead, they assume they have found the correct way to do something, and they proceed accordingly. This study assumes system designers are concerned with optimal performance.

3. Stimulus-Response Pairs

System designers are often faced with the task of choosing which of several stimuli should be used to represent a given action. For example, which of several possible icons should represent a specific user choice in a graphical user interface? Which of several possible words should represent a user choice in a speech controlled system? Which of several shapes should be manipulated at a console to produce a desired effect? If empirical testing is carried out (as it should be), the results are usually tabulated in a confusion matrix. The confusion matrix then guides the selection process.

Empirical testing of this type entails presenting test subjects with the various stimuli under consideration and tabulating the responses of the test subjects. For example, test subjects might be asked to examine a list of computer commands and their associated functions; shortly thereafter, the functions are stated one by one, and the test subjects must identify the associated function. Naturally, there will be some confusion in selecting the proper functions, but the most logical, most easily recognizable will be correctly identified most of the time. The results of all trials with all test subjects can be tabulated in confusion matrix form where the data is more easily analyzed. The analysis that follows may involve examining the commands that are most often confused and finding possible replacements for those commands.

Once the data is tabulated, however, the analyst may experience difficulty determining which are the best S-R pairs. In other words, if a subset of the S-R pairs is needed, how can the "best" subset be found? That depends partly on the analyst's definition of what "best" really means. Tools for optimally selecting subsets of stimulus-response pairs from a confusion matrix have only recently been developed (Theise, 1989). These tools have focused on the minimization of confusion within the subset and maximization of recognition. An alternative approach, appealing for its conformity with an information-theoretic framework, would be to maximize the amount of information transmitted between the stimulus and response sets. Information theory is presented in Chapter IV.

B. COMMAND, CONTROL, AND COMMUNICATIONS

1. Definition of Command and Control (C2)

Joint Chiefs of Staff Publication 1 (JCS Pub 1) defines command and control as follows:

> Command and Control: The exercise of authority and direction by a properly designated commander over assigned forces in the accomplishment of the mission. Command and control functions are performed through an arrangement of personnel, equipment, communications, facilities, and procedures which are employed by a commander in planning, directing, coordinating, and controlling forces and operations in the accomplishment of the mission. (JCS Pub 1, 1987, p. 77)

2. The Command and Control System

As equally important definition is that of a C2 system. A C2 system is:

> The facilities, equipment, communications, procedures, and personnel essential to a commander for planning, directing, and controlling operations of assigned forces pursuant to the missions assigned. (JCS Pub 1, 1987, p. 77)

A C2 system contains all the tangible elements required for command and control including communications, equipment, and procedures. These elements have very strong human factors, or human performance, ramifications. If these elements are well designed, they can be of invaluable service to the commander in his function of decision maker. The hardware involved in C2 systems is very expensive and difficult to change, as are procedures; therefore, it is imperative that the best possible systems be developed and deployed the first time to avoid the costly process of replacing ineffective or inadequate systems. (Berg, 1990, pp. 11-12)

It should also be noted at this point that, since a C2 system contains communications, by definition, the terms command and control (C2), and command, control, and communications (C3), may be used interchangeably. Typically, the term C3 is used by some to put special emphasis on communications. (Bethmann and Malloy, 1989, pp. 9-10)

3. C3 and Human Factors

It should be no small surprise that human factors plays a major role in the C2 process. The C2 process involves people interacting with machines, especially communications devices. Whenever communications takes place, there is a potential for misunderstanding or misinterpretation. This is one area where better human factors engineering or systems design would be useful. One aim of better human systems design in C3 systems is to reduce potential confusion. If some of the tools of C3 could be made more understandable, confusion would be reduced.

What are some of the tools of C3 that required human factors attention? Examples include displays on all types of electronic equipment; symbology, terminology, and physical controls such as knobs, switches, and levers. Some of these items are physical or visible while some are conceptual. However, they all require special care in their development if confusion is to be minimized.

4. C3 and Information Transfer

Another concept to consider in design is that of information and its requisite transfer. After all, there is no communications without the transfer of information. In fact, the C2 process relies heavily on information transfer. A commander cannot make decisions or give orders if he doesn't receive and transmit information in some way. Furthermore, in modern warfare, a commander must receive and transmit information at ever increasing speeds if the enemy is to be defeated.

The state of modern technology in this information age affords these ever increasing speeds, but guarantees nothing of the quality of the information being transferred. The best equipment in the world cannot turn a useless input into transferred information, but it will get there quickly and efficiently. The old adage "garbage in, garbage out" applies here.

5. Boyd's O-O-D-A Loop

As further testimony to the need for more speed and less confusion in the C2 process, many C2 experts and analysts use the work of John Boyd and his O-O-D-A loop when discussing the C2 decision making process. Several derivations of Boyd's model have been developed, but all stay basically true to the original model with slight refinements. The basic Boyd model will be used in this work.

a. *The O-O-D-A Loop*

John Boyd developed a model of the decision making process that is typically referred to as the O-O-D-A loop. The four-letter, hyphenated acronym stands for Observe, Orient, Decide, and Act. The model structure is shown in Figure 1. (Orr, 1983, p. 23-27)

The process is self explanatory. The decision maker observes the environment relative to "the problem" and the decision he faces. Next, he orients himself and the variables under his control to the situation. This involves processing and analyzing the data gathered from the observations made in the previous step. The next step requires the decision maker to make a decision, and the final step puts that decision into action. This is a very

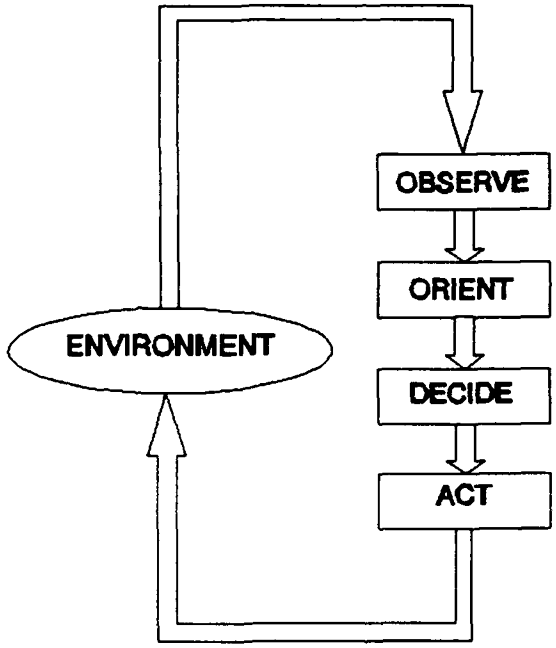

Figure 1 Boyd's O-O-D-A Loop

simplified overview of the model, but the essence of the process is all that is required here. (Orr, 1983, p. 24-30)

b. C3 and the O-O-D-A Loop

When the commander uses this process, communications must take place. The commander must receive intelligence and other information from various sources, and he must transmit his decisions and requirements to the appropriate receivers. In a combat situation, the commander must not only perform this task with little or no errors, but he must also do it quicker than the enemy can carry out their version of these same functions. Whoever can process and move through their O-O-D-A loop more quickly holds a decided advantage in a combat situation. The process is complicated by the "fog of war" which makes mistakes more likely, requiring a system with a reduced likelihood of errors.

If a system could be developed that was more efficient and effective at transferring information, the process would be improved. There are probably many steps that could be taken to reduce errors and improve system efficiency and effectiveness. One of those steps is examined here; attempting to increase information transmitted in the stimulus-response process. In this case, the commander receives a stimulus and returns an appropriate response.

This is a case where systems designers need to ensure that the system being built or redesigned uses the best possible human-system interface they can produce. One methodology available to systems designers for this purpose is operations research, including optimization techniques such as linear programming. Neither operations research nor any other method can guarantee perfection, but they can work to minimize errors, or in this case maximize information transmitted between stimulus and response. The concept of transmitted-information, as well as information theory in general, will be covered in Chapter IV.

C. OPTIMIZATION AND C3 EXAMPLES

The following examples give a feel for the need for optimal design in human interface systems. Information is a basic commodity in each of these examples; therefore, it makes sense to think of optimizing transmitted-information in these examples and other similar situations.

1. An Aircraft Example

Although not a classic C3 example, this aircraft cockpit design example contains excellent examples of potential confusion and helps introduce the idea of information transfer.

In an aircraft cockpit, there are myriad levers, buttons, switches, and displays that control the aircraft or provide information to the pilot. How does the pilot remember where everything is? How does he avoid using the wrong control for a given situation? One solution is to label everything; however, some things must become so second nature to a pilot that labels are insufficient for preventing mistakes. A better solution gives each control a specific shape enabling the pilot to feel the control, identifying it by touch. In fact, shape-coding aircraft controls is now standard practice. But if shape-coding aids discriminability between different controls, what determines the most appropriate shape for any given control? For example, if the flaps were controlled by a lever, would it make more sense to shape the gripping surface of the lever like a flap (or wing-like shape) or some other shape? In time the pilot would adapt to either one, but which would be a better a priori choice? Which control shape would "tell" the pilot more? (Kantowitz and Sorkin, 1983, 309-317)

The last question implies a transfer of information from the lever to the pilot. In fact, if there were no transfer of information, the pilot would have no reason to use the lever. In other words, if the stimulus conveys no information to the user, the user has no reason to respond to the stimulus.

2. Display Design Example

The design of displays is another excellent example of a potential source of confusion. If the display layout is not conducive to the operational environment in which it will be used, or the symbology is not well conceived, the human operators will be more likely to make mistakes when relying on the displays, or may choose not to rely on them at all if they can be avoided. Two display design examples follow.

a. *Radar Display*

An experiment was carried out at the late 1950s by Bowen, Andreassi, Truax, and Orlansky (1960) to choose an optimal set of geometric symbols for radar displays. It was believed that certain attributes were favorable such as simplicity, symmetry, and familiarity. These attributes are obviously chosen with the human operator in mind. The experiment presented subjects with various symbols, under various display conditions (noisy, distorted, blurred), with the intent of having them indicate on a score sheet which symbol they had just seen. The results were tabulated and judgements about the optimal subsets of various sizes were made. The objective, of course, was to find a set of symbols whose attributes greatly reduced the likelihood of intersymbol confusion.

Additionally, the idea of complex, auxiliary symbols was mentioned. These symbols would be made up of combinations of the basic symbol set. So, for example, if a square and a triangle each had their separate meanings, a triangle inside of a square would have yet another meaning; most likely, a hybrid meaning that would be a combination of the two separate meanings. The data for this experiment is included here as one of the test data sets called Bowen.

b. 465L System

In the late 1950s, Strategic Air Command (SAC) was developing a computer-based command and control system known as the 465L. As it turned out, users were unhappy with the system because they were required to "go from display to display to pull together the elements of the problem." Parsons, 1972, p. 349) The users felt that fewer displays that contained more complete information would be a better way to get the full situation they were attempting to assess. Here, the concept of more information from an interface device arose after users experimented with the system. How should system designers decide on the appropriate symbols to use? They could simply use the method mentioned in the previous section concerning radar displays; although, it makes sense in today's high technology environment to use mathematical tools to find the optimal set of symbols or the optimal design of a display.

3. New Global C2 Architecture

The world is changing at a rapid pace and, in an attempt to more adequately face the future, the Joint Staff conducted a study through the C2 Functional Analysis and Consolidation Review Panel (FACRP) to determine the C2 requirements for the future. The report focused on such concepts as a global C2 infrastructure capable of supporting joint and combined operations. Developing an architecture that would be interoperable with and acceptable to all concerned parties is no small task. Of particular interest to this thesis are the human factors ramifications. A global architecture means not just equipment, but policies and procedures as well. Part of the process involves agreement on terms, concepts, symbols, etc. The report mentions a requirement to transfer information via displays and interfaces. (FACRP Report, 1991, pp. 24-30) Designers should naturally desire displays and interfaces that transfer as much information as possible with the least amount of interaction or actual transmission. In other words, make the displays and interfaces as meaningful as possible so as to minimize the amount of raw data transfer. This is not a simple task considering the diversity of experience and culture in joint and combined operations. Experiments need to be conducted to decide on things such as terms, symbols, and concepts that would convey the desired meaning to all possible users. The report stresses modularity and flexibility. To achieve these goals, very careful design of the aforementioned items is required. Optimal information transfer should be a goal of system designers when developing this new global architecture.

D. OPTIMIZATION SOFTWARE

Optimization algorithms can be very sophisticated, and can require an enormous number of repetitive arithmetic calculations. Today, there are software packages available that will do all the calculations needed, and will do them very quickly. For linear programming, LINDO (Schrage, 1987) has long been one of the most widely used programs in existence. Today, LINDO is available in many forms including a PC version. LINDO required the user to completely specify the problem under consideration with objective function, constraints, and data on a case by case basis. In other words, generic models for a class of problem could not be entered for long term use. Each model had to be individually produced. Some advances to this process were made using matrix generators to generate the case specific equations rather than entering them individually.

However, matrix generators and linear programming packages are losing ground to computer-readable modeling languages. (Fourer, 1983, pp. 144-169) These software packages will take an algebraic set of expressions and generate the case specific equations for the model ready for values to be plugged in for the variables. In other words, the software program transforms algebraic form into a form that a mathematical solver program can interpret. The model produced may be a very generic model for a class of problems that is capable of reading a data file containing case specific data, additional parameters, or additional constraints.

The modeling language used in this case was the General Algebraic Modeling System (GAMS) (Brooke, Kendrick, and Meeraus, 1988). To understand the power of a model system such as GAMS consider a problem based on a 3 x 4 matrix (rows=i=3, columns=j=4). GAMS will allow an algebraic expression such as:

$$\Sigma_{i=1} x_{ij} = s_j \quad \text{for all } j$$

to be written as:

```
SUM(I, X(I,J) ) =E= S(J).
```

In turn, GAMS generates the equations:

$$x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4} = s_1$$

$$x_{2,1} + x_{2,2} + x_{2,3} + x_{2,4} = s_2$$

$$x_{3,1} + x_{3,2} + x_{3,3} + x_{3,4} = s_3$$

This is a very convenient tool, especially when the algebraic expression becomes complicated or when the expression represents a large number of possible iterations such as when the matrix in the above example becomes very large. Past linear programming methods required complete equation specification via user entry or matrix generation to produce the necessary equations suitable for solving. Additionally, these methods had data values tied directly to the equations. Modeling languages generate generic sets of equations independent of specific data values. The generic equations, or models, can then be augmented by separate data files.

GAMS is a very useful program that acts as a front-end processor for mathematical solver programs. GAMS generates equations from algebraic expressions, performs pre-solve and post-solve calculations, and provides for output data formatting. The mathematical solvers are capable of solving specific types or forms of problems and have the task of optimizing sets of equations. Some of the solvers available for use with GAMS are Zero/One Optimization Method (ZOOM) (Marsten and Singhal, 1988) for models with binary and general integer variables, Modular In-core Nonlinear Optimization System (MINOS) (Gill, Murray, Murtagh, Sanders, and Wright, 1988) for nonlinear and general optimization models with continuous variables, and XA (Sunset Software Technology, 1987) a very fast and powerful integer program solver. For a more elaborate description of these software packages, see *GAMS: A User's Guide* by Brooke, Kendrick, and Meeraus (1988).

III. THE CONFUSION APPROACH TO OPTIMIZATION

One successful attempt that has been made at optimization in human factors engineering was the work on minimizing confusion done by Theise (1989) that was mentioned in Chapter I. Theise proposed that if confusion between various stimuli could be minimized, mistakes would be much less likely. This method relies on confusion matrices and binary integer programming. Confusion matrices were briefly discussed in the Introduction. A brief review of confusion matrices and their use is presented in this chapter.

A. THE CONFUSION MATRIX

Analysis in the area of discriminability has been going on for years, taking many evolutionary turns. The shape-coding of aircraft controls comes from early empirical research in the area of discriminability and confusion. Empirical analysis usually involved experiments where subjects were presented with stimuli and prompted for a response. The results were tabulated in a confusion matrix where recognition between a stimulus and its proper response is tabulated on the main diagonal, and confusion between stimuli and responses is tabulated on the off-diagonal. A simple example of a confusion matrix was presented in Table 1.

In early analysis, picking subsets of S-R pairs from a matrix was usually done by simply examining the matrix and selecting the pairs that appeared to have little interaction with each other--'eyeballing it.' Eyeballing it can be rather easy if the confusion matrix is small and sparse but becomes increasingly difficult as the matrix becomes larger or more dense.

B. CLUSTER ANALYSIS

As this area of study grew, a more scientific process called cluster analysis was applied. Cluster analysis entails the formation of clusters of S-R pairs based on similarity. The objective is to ensure a high degree of confusion within clusters but a relatively low degree of confusion between clusters. Once the clusters have been formed, subsets can be formed by selecting S-R pairs from different clusters. Because the clusters have a low degree of intercluster confusion, selecting from different clusters should imply low overall confusion within the selected subset, but this is not always the case. One weakness of some types of cluster analysis is the inconsistency in the composition and interpretation of the clusters from analyst to analyst. Although still in wide use today, it is not a completely deterministic method, and therefore lacks optimality. Like 'eyeballing it,' cluster analysis becomes more difficult as matrix density increases. A full discussion of cluster analysis including its use on confusion matrices can be found in *Cluster Analysis for Researchers* by Romesburg (1984). A detailed description of clustering algorithms can be found in *Algorithms for Clustering Data* by Jain and Dubes (1988).

C. THEISE'S CONFUSION/RECOGNITION MODELS

Recently, Theise (1989) developed models using binary integer programming to select subsets having minimum total confusion.

1. Moore's Pushbutton Data

The primary data used by Theise in his presentation was from T.G. Moore's (1974) research in attempting to find an optimal set of pushbuttons for the British postal system. Moore published his findings in an article titled "Tactile and Kinaesthetic Aspects of Pushbuttons" in *Applied Ergonomics*, 1974. Moore's method of analysis was a form of cluster analysis known as McQuitty analysis (McQuitty, 1957). Since the data set on pushbuttons used by Moore in his research is relatively large (25 pushbuttons in the original set), it will also be used as an example in this paper. Additionally, the pushbutton data was used in two previous optimality studies so it provides an opportunity for comparison.

Figure 2 shows the 25 pushbuttons that were included in Moore's initial set. Table 1 shows the confusion matrix resulting from a test Moore conducted to determine whether tactile aspects of the pushbuttons allowed for easy distinction between the various buttons. This confusion matrix provides for the data to be used later in the Transmitted-Information Model.

The objective of Moore's research was to select six pushbuttons that would allow operators in the sorting department of the British Postal System to be able to operate the sorting machine without actually looking at the pushbuttons. Six pushbuttons with distinctive tactile aspects were needed. Moore's research resulted in the selection of pushbuttons 1, 4, 21, 22, 23, and 24. This will be compared to the selections arrived at using the Confusion/Recognition Model and the Transmitted-Information Model developed in this paper.

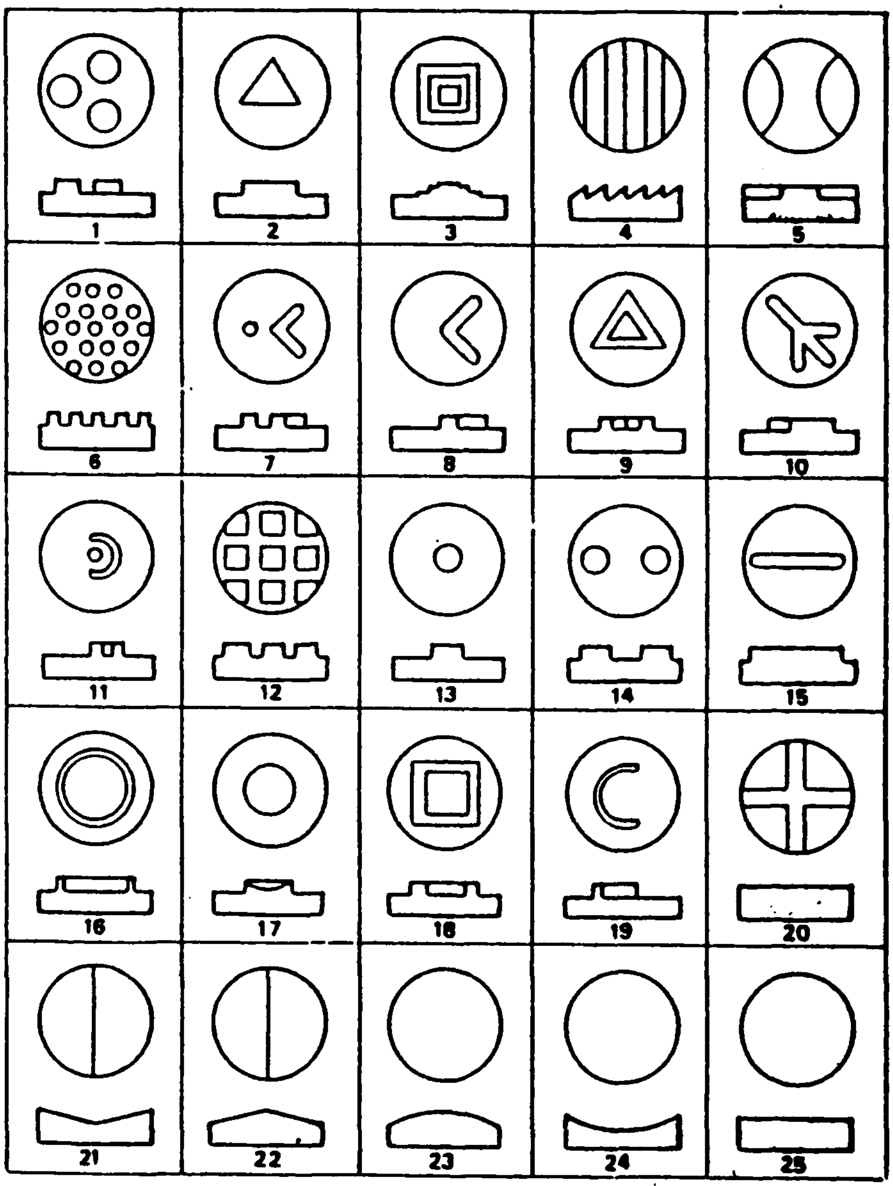

Figure 2 Pushbuttons Tested by Moore

TABLE 2 CONFUSION MATRIX FROM MOORE'S EXPERIMENT

		RESPONSES ON PUSHBUTTONS																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	U
	1	65	0	0	0	0	1	0	0	0	0	0	7	1	6	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	71	0	0	3	0	0	1	0	0	1	0	3	0	1	0	0	0	0	0	0	0	0	0	0	0
	3	0	2	47	0	2	0	3	0	0	3	5	0	8	0	0	1	1	0	0	1	0	0	0	0	0	7
	4	0	0	0	71	0	3	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	10	0	0	57	0	0	0	1	0	0	1	0	0	3	0	0	0	0	7	0	0	0	0	1	0
S	6	1	0	0	1	0	45	0	1	0	1	1	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	7	2	1	2	0	0	0	58	4	2	7	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0
I	8	0	1	1	0	0	0	0	60	0	0	0	0	0	0	2	1	0	0	15	0	0	0	0	0	0	0
M	9	0	1	0	0	1	0	2	2	63	0	1	1	0	0	0	0	0	8	1	0	0	0	0	0	0	0
U	10	0	0	0	0	0	3	16	0	1	39	10	2	0	0	0	1	0	4	0	4	0	0	0	0	0	0
L	11	1	2	4	0	0	4	14	0	0	4	37	0	1	7	0	0	0	1	0	0	0	0	0	0	0	5
I	12	1	0	0	8	0	22	1	0	0	0	1	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	13	0	4	2	0	0	0	1	0	0	0	0	0	72	1	0	0	0	0	0	0	0	0	0	0	0	0
	14	7	0	0	0	0	1	1	0	0	0	0	0	0	71	0	0	0	0	0	0	0	0	0	0	0	0
	15	0	2	0	0	4	0	0	3	0	1	0	0	0	0	69	0	0	0	0	1	0	0	0	0	0	0
	16	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	50	9	10	1	0	0	0	0	1	6	0
	17	0	1	0	0	0	1	0	0	2	1	0	0	0	0	0	4	65	2	0	0	0	0	0	1	3	0
	18	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	7	0	60	1	0	0	0	0	0	0	0
	19	0	0	0	0	0	0	0	4	21	0	0	0	0	0	0	0	0	1	54	0	0	0	0	0	0	0
	20	0	0	1	0	4	0	1	0	2	3	1	0	0	0	1	0	0	0	0	67	0	0	0	0	0	0
	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	79	0	0	0	0	0
	22	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	75	1	0	1	0
	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	79	0	0	0
	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	75	1	0
	25	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1	73	0

2. The Confusion/Recognition Models

Theise (1989) developed four models with the underlying objective of minimizing confusion. The models select optimal subsets of S-R pairs with minor variations from one model to the next--one model bases selection strictly on minimizing confusion while another attempts to maximize recognition subsequent to minimizing confusion. These models exhibit the deterministic nature lacking in previous methods of subset selection and they may find wide use as their utility is uncovered by system designers and analysts. The primary interest here will be on Theise's third model, aimed at minimizing confusion while maximizing recognition. (Theise, 1989, pp.298-300) Theise called this model *The Maximum Total Recognition Given Minimum Total Confusion Problem*, in this paper it will be referred to as the *Confusion/Recognition Model*.

a. *The Minimal Confusion Model--Model 1*

The minimal confusion model (Model 1) is actually quite simple. The objective function is simply a summation of all of the off-diagonal values in the selected subset with a constraint ensuring the selected subset size is correct. These optimization equations are shown below. Note the u_i variable is included to handle cases where no response was given to a test stimulus. (Theise, 1989, pp. 297-298)

$$\text{Minimize} \quad \Sigma_{i=1}\Sigma_{j=i+1}C_{ij}x_ix_j + \Sigma_{i=1}u_ix_i$$

$$\text{Subject to} \quad \Sigma_{i=1}x_i = s$$

$$x_i \text{ binary}$$

An additional constraint is required here due to the limitations of the software package. The problem lies in the inability of the mathematical solver to handle binary integer variables and nonlinearities simultaneously. This is present in the objective function in the form of the term x_ix_j where the product of two binary integer variable is required to select each confusion value being summed in the objective function. Each value in the matrix is identified by a "row" variable and a "column" variable. Since this situation cannot be handled by the solver, an alternative method of identifying the individual confusion values is needed. Theise solved this problem using a well known linearization technique wherein the binary integer variable y_{ij} is substituted for the x_ix_j term and the following linear constraints are added. (Phillips, Ravindran, and Solberg, 1987, pp. 190-191)

$$\left.\begin{array}{r} x_i + x_j - y_{ij} \leq 1 \\ -x_i - x_j + 2y_{ij} \leq 0 \end{array}\right\} \text{ for all } C_{ij} > 0;\ i \neq j$$

The first constraint ensures that when both x_i and x_j are equal to one, y_{ij} will be forced to equal one to maintain the inequality. This ensures that the proper confusion values are included in the summation. The second constraint forces y_{ij} to equal zero under all other circumstances such as when only one of x_i or x_j is equal to one. Close examination reveals that only the first of these new constraints is needed. Since x_i, x_j and y_{ij} are all binary variables, they can only have the values 0 or 1; additionally, since the objective is to minimize, the solver will try to make these values 0 wherever possible. If either x_i or x_j is 0, y_{ij} will be 0 due to the objective function. If x_i and x_j are both 1, y_{ij} will be forced to be 1 and the confusion value will be included. Consequently, the second new constraint would be redundant. This confusion model will now sum only the off-diagonal values of confusion for the S-R pairs included in the selected subset.

b. Confusion/Recognition Model--Model 3

Model 3 seeks to ensure not just minimum confusion, but also maximizes recognition as a secondary consideration. In other words, minimize confusion first, then, given the minimum confusion, maximize recognition.

The additional notation required for this model includes a variable d^+ which measures the positive deviation in total confusion from a specified threshold t. The threshold is typically preset to a value of zero. Furthermore, a large positive constant was required to be used as a penalty cost for deviating from the confusion threshold. The constant M was defined, for convenience, as the sum of all the confusion values in the matrix as shown in the following equation.

$$M = \Sigma_{i=1}\Sigma_{j=1}C_{ij} + \Sigma_{i=1}u_i$$

The entire model is as follows:

$$\text{Maximize} \quad \Sigma_{i=1} c_{ii}x_i - Md^+$$

$$\text{Subject to} \quad \Sigma_{i=1}\Sigma_{j=i+1}C_{ij}y_{ij} + \Sigma_{i=1}u_ix_i - d^+ \leq t$$

$$\Sigma_{i=1}x_i = s$$

$$x_i + x_j - y_{ij} \leq 1 \quad \text{for all } C_{ij} > 0;\ i \neq j$$

$$x_i \text{ binary}$$

The objective function sums the diagonal values of the selected subset. This, of course, represents recognition. The value subtracted from this sum is a penalty cost for exceeding the threshold value of confusion set by the first listed constraint. Since M is a large value, a large penalty is paid for exceeding the threshold value; in fact, in the objective function, the term ($- Md^+$) is more influential than the sum of the recognition values. The first constraint ensures that the sum of the off-diagonal values (confusion values) in the selected subset is minimized by ensuring this sum is less than the predetermined threshold value. If this is not the case, the value of d^+ increases causing a large penalty to be paid in the objective function. Therefore, the model will always try to minimize confusion first, and maximize recognition second. The other two constraints operate exactly as they had in Model 1.

Note that in these models the only confusion values above the main diagonal are summed. This is because the confusion matrix is triangularized. This could easily be done by the model by changing the first constraint to the following:

$$\Sigma_{i=1}\Sigma_{j=i+1}(C_{ij} + c_{ji})y_{ij} + \Sigma_{i=1}u_i x_i - d^+ \leq t$$

This modification has the effect of triangularizing the matrix.

c. Confusion/Recognition Model Results

For Moore's data, the Confusion/Recognition Model selected a subset of pushbuttons 2, 4, 14, 20, 21, and 23 with a total value of zero for confusion which, incidentally, is the lowest value possible since negative confusion values are undefined. A value of 438 was found for recognition. If confusion and recognition were totaled in the same way for the subset Moore selected using cluster analysis, the confusion value would be five and the recognition value would be 444. The confusion value is not very large but there are actually many possible subsets with zero total confusion. Also note that the recognition is higher in Moore's subset, but this comes at the expense of the higher confusion value. (Theise, 1989, p. 302)

Based on confusion/recognition it appears as though Moore failed to select the optimal subset. If optimality were based on just confusion, his choice is still not optimal.

However, if recognition alone were used to select the optimal subset, Moore's selection has a higher value than the subset selected by the Confusion/Recognition Model. But, Moore's subset was not optimal in terms of recognition either. In fact, the maximum recognition subset contains pushbuttons 13, 21, 22, 23, 24, and 25, and has a recognition value of 453. Unfortunately, this subset also has a confusion value of 13. The primary consideration here is the question of what is the "best" subset or what is the best method for selecting the "optimal" subset. The basic premise of the Confusion/Recognition Model appears sound. After all, minimizing confusion is a very desirable action in a human-system interface. Furthermore, once confusion has been minimized, selecting what is most easily recognized is also desirable. It is important to remember at this point that any model is only as good as the data applied to it and the experiment that produced the data.

IV. INFORMATION THEORY

A. INTRODUCTION

Another analytic approach to the problem comes from the realm of information theory. It has been demonstrated that given a confusion matrix, the total amount of information transmitted by all S-R pairs in the matrix can be calculated using information theory and basic set theory (Kantowitz and Sorkin, 1983, pp. 142-143; Garner, 1962, pp. 19-58). The prospect of marrying the binary integer programming approach to information theory is appealing for its conformity to the information theoretic framework; a well accepted body of knowledge exists in areas of study such as human factors, communications engineering, and statistics and experimental design.

B. OVERVIEW OF INFORMATION THEORY

The theory and notation in this section is taken primarily from Garner (1962). Additional notation and theory comes from Kantowitz and Sorkin (1983).

1. Information Theory Background

Information theory is derived from communications theory and is motivated by a desire to quantify information as a measurable commodity. By definition, when communications occurs, information must be transmitted. Note that, regardless of how information is measured, the measurement tells nothing of the value of the information. Value is determined by the recipient or user of the information. Before the amount of information can be explored, the basic properties of information must be examined.

> Information exists in a message or communication only if there is an *a priori* uncertainty about what the message will be. (Garner, 1962, p. 3)

In other words, if the receiver is already aware of the facts contained within the message, then no information has been received. If it is raining outside and the receiver is gazing out the window, he will learn nothing if someone tells him it is raining. He has, therefore, received no information because he has no uncertainty about whether it is raining or not. However, if he is told that the total rainfall over the past hour was 0.15 inches, information has been transmitted because he was not previously aware of the amount of rainfall--he was uncertain.

Furthermore, the amount of transmitted-information is determined by the amount of uncertainty "...or, more exactly, it is determined by the amount by which uncertainty has been reduced." (Garner, 1962, p. 3) An example illustrates this point. Consider a fair coin that is to be tossed. Before the coin is tossed, there is no a priori knowledge of the outcome since the outcome of a fair coin toss is equally likely to heads as tails i.e., we are completely uncertain. After the coin has been tossed, the outcome is known, the uncertainty has been removed, and information has been gained. If there were to be multiple tosses of the coin, there would be that much more uncertainty about the overall outcome--the total number of heads for example. One toss of a fair coin results in the resolution of a situation that had two possible outcomes, while two tosses of a fair coin has four possible outcomes, and three tosses has eight possible outcomes. Specifying information in this way is cumbersome, so a simpler method was developed. The measure must "satisfy the two conditions that (a) it is monotonically related to the number of possible outcomes and, (b) each successive event adds the same amount of uncertainty and thus makes available the same amount of information." (Garner, 1962, p. 4) This a logarithmic relationship and for reasons of proportionality, the base was chosen to be two. The following equation gives a basic measurement of information:

(1) $$U = \log_2 m$$

where U is the measure of uncertainty and, therefore, information, and m is the number of possible outcomes. The unit of measure is the bit, commonly used in communications and computer technology. So, if a fair coin is tossed, one bit of information has been gained because one bit of uncertainty has been resolved. Likewise, if eight coin tosses are made eight bits of information are gained. (Note that for eight coin tosses, there are 256 possible outcomes and $U = \log_2(256) = 8$.)

2. Developing a Concept of Information Measurement

The next step in developing the information measurement concept is to extend the process to situations where the possible outcomes are expressed as probabilities rather than a strict enumeration. The probability of occurrence of any event is the reciprocal of the number of possible outcomes, so equation (1) becomes:

$$U = \log_2(1/p(x)) = -\log_2 p(x) \tag{2}$$

where p(x) is the probability of the outcome of x.

To sum up the total information contained over a long term and over several categories of events, a weighted average must be taken. The equation which expresses the average uncertainty associated with a discrete probability distribution is given by:

$$U(x) = -\Sigma p(x)\log_2 p(x). \tag{3}$$

This concept can easily be extended to two variables x and y. In this case, the concern is with the joint occurrence of events x and y. The uncertainty involved in this joint occurrence is found by:

$$U(x,y) = -\Sigma p(x,y)\log_2 p(x,y). \tag{4}$$

This is referred to as the *joint uncertainty*, and p(x,y) is the joint probability, or probability of x and y occurring. Typically, the variables, x and y, are correlated;

consequently, $p(x,y) \neq p(x)p(y)$. The uncertainty that would exist if x and y were not correlated is a value that has utility in this development, so it is presented here. It is referred to as *maximum joint uncertainty* because it is the highest level of uncertainty possible with the given values of $p(x)$ and $p(y)$.

$$(5) \qquad U_{max}(x,y) = -\Sigma P(x,y)\log_2 P(x,y)$$

The difference between maximum joint uncertainty and joint uncertainty is called *contingent uncertainty* (the uncertainty contingent on the correlation of the variables) and is represented by $U(x:y)$.

$$(6) \qquad U(x:y) = U_{max}(x,y) - U(x,y)$$

$U(x:y)$ will also be referred to as *INFO* in this paper. As correlation between x and y increases the value of joint uncertainty decreases, so contingent uncertainty would increase thus illustrating that it represents the amount by which uncertainty is reduced by the correlation. In other words, if joint uncertainty is maximum (no correlation), then contingent uncertainty is zero--uncertainty hasn't been reduced at all. Conversely, if joint uncertainty is minimum (high degree of correlation), then contingent uncertainty is high--uncertainty has been reduced a great deal by correlation. According to Garner, "one of the most common uses of the contingent uncertainty is as a measure of information transmission." (Garner, 1962, p. 63)

C. INFORMATION MEASUREMENT EXAMPLE

To illustrate the use of information theory in quantifying the available information contained within the stimulus-response pairs in a confusion matrix, a sample set of calculations is presented here. The data used comes from the simple confusion matrix presented earlier in Table 1. (Clarke, 1957, pp. 715-720)

The first calculation is to determine the joint uncertainty, U(x,y), using equation (4); however, to find the joint uncertainty, the probability of each cell, the $\log_2$ of that probability, the negative of the product of these two values, and, finally, the sum of these products are needed. In fact, this sum is the joint uncertainty. The values shown in Table 3 are in the form $-p(x,y)\log_2 p(x,y)$. Note that if a cell had a zero probability, it would not require any further calculation; the $p(x,y)\log_2 p(x,y)$ is evaluated as zero. The joint uncertainty is the sum of all the values in Table 3. This sum, U(x,y), is 4.5436.

The next step is to calculate the maximum joint uncertainty, $U_{max}(x,y)$ equation (5). To find this value, similar calculations to those done for joint uncertainty are required, but for maximum joint uncertainty, each row and column are treated individually. The pertinent row and column values required for the maximum joint uncertainty calculation are shown in Table 4. As Table 4 illustrates, the maximum joint uncertainty, $U_{max}(x,y)$, is 5.1483.

TABLE 3 CALCULATING JOINT UNCERTAINTY

	pa	ta	ka	fa	θa	sa
pa	0.2625	0.1868	0.1407	0.1184	0.0557	0.0216
ta	0.2127	0.2251	0.1820	0.0870	0.0529	0.0329
ka	0.1681	0.2764	0.1858	0.0308	0.0638	0.0403
fa	0.0962	0.0216	0.0216	0.3503	0.1413	0.0576
θa	0.0647	0.0576	0.0482	0.2232	0.2347	0.1618
sa	0.0179	0.0814	0.0576	0.0433	0.2073	0.3137

TABLE 4 CALCULATING MAXIMUM JOINT UNCERTAINTY

Stimulus/ Response	p(x)	$-p(x)\ Log_2 p(x)$
Row pa	0.1667	0.4308
Row ta	0.1667	0.4308
Row ka	0.1667	0.4308
Row fa	0.1667	0.4308
Row θa	0.1667	0.4308
Row sa	0.1667	0.4308
Column pa	0.1788	0.4441
Column ta	0.1907	0.4559
Column ka	0.1233	0.3724
Column fa	0.2077	0.4709
Column θa	0.1558	0.4179
Column sa	0.1437	0.4022
	Total:	5.1483

Information transmitted, also called contingent uncertainty, $U(x:y)$, is found by evaluating equation (6). Therefore, information transmitted by the six S-R pairs evaluated is:

$$U(x:y) = U_{max}(x,y) - U(x,y) = 5.1483 - 4.5436 = 0.6047 \text{ bits}$$

V. MAXIMAL INFORMATION SUBSETS

A. THE CONCEPT OF MAXIMAL INFORMATION

Using the calculations from the previous chapter, information transmitted could be calculated for any number of S-R pairs. For example, in the sample calculations at the end of Chapter IV, all of the S-R pairs were used to find information transmitted. If only two of the six S-R pairs were required for a specific application, the question is which two should be used. From the perspective of transmitted-information, it makes sense to use the two S-R pairs that transmit more information combined than any other two S-R pairs combined. Using the same data from the previous example, the following table shows the transmitted-information (the U(x:y) column) by all possible combinations of two S-R pairs.

From the data in Table 5, it should be obvious that the choice of S-R pairs *pa* & *sa* results in the maximal transmitted-information for a subset size of two. If the objective is to maximize transmitted-information using only two of the S-R pairs, these two S-R pairs should be selected since, together, they transmit 0.8035 bits of information.

TABLE 5 TRANSMITTED-INFORMATION FOR SUBSETS OF SIZE TWO

S-R Pairs	U(x:y)
pa & ta	0.0159
pa & ka	0.0467
pa & fa	0.3115
pa & θa	0.4547
pa & sa	0.8035
ta & ka	0.0036
ta & fa	0.5186
ta & θa	0.4534
ta & sa	0.4924
ka & fa	0.6135
ka & θa	0.3964
ka & sa	0.4699
fa & θa	0.0826
fa & sa	0.6450
θa & sa	0.0595

Obviously, this method of determining the optimal subset for transmitted-information would become extremely tedious if the number of original S-R pairs became much bigger than four; a very real probability. The number of subsets of size s selected from a group of size n that must be evaluated to perform a complete enumeration is found using the well known formula for combinations:

$$\frac{n!}{(n-s)!\,s!}$$

For example, if the original number of S-R pairs is ten (n=10) and a subset of five pairs is desired (s=5), then 252 subsets must be investigated since there are 252 subsets of size five when selecting from a group of ten. Furthermore, the Moore data set (25 S-R pairs) has 177,100 subsets of size six which

Moore was attempting to select. Performing these calculations by hand would be, as previously stated, extremely tedious and time consuming. With the computer technology available today, there should be an easier method. The method of interest here not only lets computer software calculate the information values, but also allows the software to select the optimal subset. This is possible using a software package such as GAMS. The next section discusses the development of a GAMS model for the purpose of selecting maximal transmitted-information subsets.

B. DEVELOPING A MODEL FOR MAXIMAL TRANSMISSION OF INFORMATION

The confusion matrix form constituted the guiding element in the development of the model. Using the values from this confusion matrix, equation (4) is transformed into:

(7) $$U(s,r) = -\Sigma_i\Sigma_j[(C_{ij}/T)\log_2(C_{ij}/T)]$$

where $T = \Sigma_i\Sigma_jC_{ij}$. Equation (5) is transformed into:

(8) $$U_{max}(s,r) = -\Sigma_iS_i\log_2S_i - \Sigma_iR_i\log_2R_i$$

where S_i is the probability of a stimulus occurring in row i and R_j is the probability of a response occurring in column j. (Note: s and r will be used in place of x and y as arguments in model equations from this point on while x and y will be used to represent binary or "switch" variables.)

This leads to a restatement of equation (6) as

(9) $$INFO = U(s:r) = -\Sigma_iS_i\log_2S_i - \Sigma_iR_i\log_2R_i - \Sigma_i\Sigma_j[(C_{ij}/T)\log_2(C_{ij}/T)]$$

The model developed must be capable of selecting a subset of these S-R pairs so as to maximize U(s:r). The simplest way to use binary variables in a case like this is to multiply each occurrence of a C_{ij} by a binary variable. Actually, this case requires each C_{ij} to be multiplied by two binary variables, x_i and x_j, because each value of C_{ij} selected must be selected by a stimulus variable and a response variable; therefore, each occurrence of C_{ij} is multiplied by $x_i x_j$ to control its inclusion or exclusion in the selected subset. So, if in Figure 1, S-R pairs 1 and 3 are selected, then all C_{ij} contained in rows 1 and 3 that are also contained in columns 1 and 3 will be used in the calculations. These values are C_{11}, C_{13}, C_{31}, and C_{33}, and each of these values needs to be multiplied by $x_1 x_3$, where both x_1 and x_3 are equal to one and all other $x_i x_j$ pairs are equal to zero. If this is true, then only the desired values of C_{ij} will be included in the selected subset.

So far, the development of the model has been quite simple. However, on closer examination, equation (9) now contains binary variables and nonlinear terms, a condition no solver can currently handle. In fact, there are nonlinearities in each of the three terms in equation (9) causing a complete failure of the model as developed thus far.

Approximation is the next logical step. If stimuli are assumed to be equiprobable, and subsequently responses are also considered equiprobable, then the U_{max} term can be considered constant, and can thus be removed from the model. Is this a reasonable approximation? Perhaps. The original premise in information theory was that this is the maximum possible uncertainty given the row and column probabilities, so although U_{max} is not, in fact, a constant, it is not completely unreasonable to approximate this value as a constant for a given subset size. Therefore, U_{max} will be considered constant for this model and empirical testing will determine if the approximation is reasonable or not. Since the objective of the model is to find an optimal subset, the quantity used to determine optimality is not as vital as the actual determination of the optimal subset. Therefore, rather than calculate a constant to be used in place of U_{max}, U_{max} will simply be dropped from the equation. Information transmitted by the selected subset can be found precisely using post-solve calculations in the GAMS model.

The approximation reduces the equation to:

(10) $$INFO = -\Sigma_i\Sigma_j[(x_ix_jC_{ij}/T)\log_2(x_ix_jC_{ij}/T)]$$

Notice that this equation is actually a form of equation (4). In other words, the model has been reduced to the joint uncertainty equation. If equation (6) is examined, it is apparent that in order to maximize U(s:r) (information transmitted), U(s,r) (joint uncertainty), must be minimized, assuming U_{max} is constant. A problem still exists in this model because it is still nonlinear and contains binary variables. Nonlinearities exist in the log term (taking the log of a binary variable) and also in the $x_i x_j C_{ij}/T$ term because T contains binary variables also. Recall, $T = \Sigma_i \Sigma_j C_{ij}$ but all C_{ij} terms must be multiplied by binary variables, so division of binary variables also exists. In fact, the product, $x_i x_j$, is another source of nonlinearity. These problems will be dealt with one at a time.

Using the same assumptions used to remove the U_{max}, the T term can be approximated by using a scaled version of the total for the entire set rather than the true total for the selected subset. To produce a value that is properly scaled the T term is scaled by the value s/n where s is the desired subset size and n is the size of the original set. As with the previous approximation, this approximation assumes the matrix is made up of equiprobable elements.

The equation has now been reduced to:

$$(11) \qquad U(s:r) = -\sum_i \sum_j \frac{C_{ij}x_i x_j}{\frac{s}{n}T} Log_2 \frac{C_{ij}x_i x_j}{\frac{s}{n}T}$$

Now, the argument of the log term can be treated as a constant term in the summation and the binary variables can be moved outside of the log term. This step allows the log term to be evaluated as a pre-solve calculation. In fact, when the binary variables are removed from the argument of the log term, the entire equation becomes the summation of constants that are chosen by binary variables. The confusion matrix can, therefore, be converted to a matrix of probabilities further transformed by the $\log_2$. In the model these values are represented by the parameter LP(I,J) and the model is now reduced to

(12) $$INFO = \Sigma_{ij} LP_{ij} x_i x_j$$

where the LP_{ij} terms are determined by

(13) $$LP_{ij} = p_{ij} \log_2(1/p_{ij}) \quad \text{all } i,j$$

and each p_{ij} term is determined by

(14) $$p_{ij} = nC_{ij}/sT \quad \text{all } i,j$$

There is still a problem with the product $x_i x_j$ but that is easily rectified. Rather than multiply the terms x_i and x_j, a new term, y_{ij}, is introduced. The relationship between y_{ij} and the x terms is given in the following linear equation which is included as part of the GAMS model

(15) $$x_i + x_j - y_{ij} \leq 1 \quad \text{for all } C_{ij} > 0;\ i \neq j$$

where x_i and x_j are binary variables. Because the goal is to minimize the objective function, INFO, y_{ij} will be zero whenever possible. If a S-R pair is selected, the value of y_{ij} will be forced to a value of one by equation (15). Since these conditions exist, y_{ij} doesn't have to be a binary variable, it merely needs to be limited to positive values. To make the solver's job easier, it is best to limit the number of binary variables as much as possible.

To further aid the solver in its calculations, the matrix was triangularized in the objective function. This was achieved by selecting the main diagonal values, LP_{ii}, then adding the values of LP_{ij} and LP_{ji}. Neither of these latter values would ever appear in solution exclusive of the other so they need not be treated separately. This also allows the y_{ij} values, and subsequently the x_i and x_j values, to be limited to only those where $i \leq j$, i.e., the matrix is upper triangularized. So, an additional group of variables was avoided. The fewer variables in the model, the easier time the solver will have in optimizing.

Subset size desired was controlled by the following equation also included in the model

(16) $$\Sigma_i x_i = S$$

where x_i is one if S-R pair i is included in the subset, and zero otherwise.

A further embellishment was to place the model in a loop so all subset sizes could be examined for any given set of data using only one GAMS run. Some sample data sets are included with this report as are the associated GAMS output data listings. The data set, a separate file called by the model using an INCLUDE statement, shows the run index starting at RUN02 rather than RUN01. This convention was used to simplify data analysis--run number equals subset size.

The final addition to the model was the set of post-solve calculations which calculate the actual information transmitted by the selected subset. The calculations were included because the model was designed to minimize a value that didn't accurately represent information transmitted due to approximations. The actual values of information transmitted would become useful in a comparison to the known optimal values that were empirically calculated during the analysis that took place after the model was developed and run. An additional post-solve calculation was included to show the values of confusion and recognition for the selected subset. These calculations were taken from the Confusion/Recognition Model and were included for use in comparison and evaluation of model performance in the analysis chapter. The entire model, with a sample data file, is included in Appendix A.

C. RUNNING THE MODEL

The model was run on 17 data sets. Most data sets contained ten or less stimuli; one contained 20, and one contained, 25. The Moore and Clarke confusion matrices were shown in Tables 1 and 2. The remaining confusion matrices are shown in Appendix B.

The solver had no trouble at all with the 15 smaller size data sets including the Bowen data set (20 S-R pairs); however, on the Moore data set (25 S-R pairs), the solver began to bog down at subsets of size 11. Up through size 10, the solver was reasonably quick, but above this level, the number of branch and bound iterations used by the solver exceeded 25,000 causing excessive time for solution. The model was modified to allow for more iterations and more solution time. Eventually, a more powerful solver called XA was made available in the operations research computer lab. Solution time with the XA solver was never a problem. The longest solution times were between 15 and 20 minutes for subsets of size 12, 13 and 14 for the Moore data set. The XA solver never failed to return a solution. The output data from the Transmitted-Information Model can be seen, along with data from the other models discussed in Chapter VI, in tabular and graphical forms in Appendix E.

VI. ANALYSIS OF RESULTS

A. DILEMMA: HOW TO ANALYZE THE DATA

One of the problems with collecting and collating data is finding a basis for comparison. Since the model attempts to identify the optimal subsets of size s from a set of size n, it would be very helpful to know what the optimal subsets are. First of all, when discussing human performance or human-system interface, is there a truly optimal answer? That depends on how optimal is defined for the situation. In this work, optimal is considered to be the best analytical answer (subset) given the data set. This assumes the data collection experiment was properly conducted without bias. Given the data, the optimal subset will then depend on the objective function used to gauge optimality. Theise used confusion and/or recognition. The measure of interest in this work is transmitted-information. To accomplish a comprehensive analysis, the results of the information model were examined with respect to the optimal transmitted-information value and with the optimal subsets selected by the Confusion/Recognition Model.

1. The Optimal Value of Transmitted-Information

If the optimal transmitted-information level for a given subset size is not known, how can the information model be evaluated? It was decided that an exhaustive enumeration would be attempted to find the optimal transmitted-information value, and the corresponding subset, for each subset size in each data set. The enumeration was carried out by a computer program that was written in Turbo Pascal (Borland International, 1987). The complete Turbo Pascal program listing is included in Appendix C with a sample input data file. This routine will be referred to as the enumeration scheme.

The program had to be capable of calculating the value of information transmitted by each possible combination of S-R pairs for each subset size. A literature search turned up a Pascal procedure designed specifically for the purpose of complete enumeration of a combinatorial problem. The recursive procedure shows up in the listing in Appendix B as the procedure called COMBS and is credited to Rohl (1983).

The program simply calculates the information transmitted by each possible combination of a given size and saves the five largest values, with the associated subset, in an array. The highest output value for each subset size (the optimal value of transmitted-information) and the corresponding subset chosen by the enumeration scheme are shown in the tables and graphs in Appendix E.

Initially, there were problems encountered when trying to run the enumeration scheme on the Moore data set. The program had to process as many as 5,200,300 combinations for both subsets of size 12 and 13. The solution time would have exceeded two weeks on the personal computer that was initially used (an Intel 80386-based 33MHz personal computer with math coprocessor). A more powerful Intel 80486-based personal computer was eventually used and provided an optimal subset for all subset sizes in less than 48 hours.

2. The Optimal Value of Confusion/Recognition

In addition to the optimal values returned by the enumeration scheme, the subsets selected by the Transmitted-Information Model are compared to the subsets selected by the Confusion/Recognition Model. In order to conveniently use the Confusion/Recognition Model, it had to be modified to accept various data sets. The model was put into a form nearly identical to the Transmitted-Information Model. Additionally, post-solve calculations were added to allow for simple model comparisons. The modified version of the Confusion/Recognition Model is included in Appendix D with a sample input data file.

B. AN EXAMINATION OF THE DATA

The primary emphasis in this data analysis will be on the numbers: information transmitted and confusion/recognition. Since these numbers are reflective of the subsets selected, the subsets selected will only be discussed when necessary. Note that the tables in Appendix E include the output data from all three models for comparison. Also included in the tables are the selected subsets for each data set and size.

1. Information Transmitted

The tables showing the values of information transmitted show the value from the enumeration scheme in the left column since it is the known optimal value. The next column shows the value from the Transmitted-Information Model (the model of primary interest), and the final column shows the post-solve value from the Confusion/Recognition Model.

A thorough examination of the information transmitted tables reveals a couple of trends. First, the value from the enumeration scheme is always the largest value whether it is singularly large, or equally as large as the value for one of the other models. This was expected since the enumeration scheme was designed to return the optimal value. Next, the Transmitted-Information Model returned a higher information transmitted value than the Confusion/Recognition Model in only 25 cases (there are a total of 149 cases). The Confusion/Recognition Model returned a higher information transmitted value than the Transmitted-Information Model in 30 cases. In all other cases, these two models returned the same value. In 80 cases, all three models returned the same value; consequently, the enumeration scheme returned a higher value than both the Confusion/Recognition Model and Transmitted-Information Model in the 69 remaining cases.

Lastly, in most of the cases where these models returned different values, the values were not significantly different from the standpoint of absolute numbers. Typically, the amount of deviation between values was less than ten percent; however, there were several cases where the difference was greater with values as high as 25% relative difference. The significance of the difference between the values is up to the individual user and the associated application. For some users, the graphs in Appendix E give a better visual presentation of the potential significance between results returned the three models.

2. Confusion/Recognition

The tables in Appendix E also include the confusion/recognition values for the optimal subsets selected by each model or scheme. The confusion/recognition values listed for the Confusion/Recognition Model are the optimal solution results from the model. The confusion/recognition values listed for the Transmitted-Information Model and the enumeration scheme are from post-solve calculations based on the maximal transmitted-information subsets selected by these models. The data is listed in the form:

confusion recognition.

Recall that the primary objective is to minimize confusion, and the secondary objective is to maximize recognition. This data is also shown in graphical form in Appendix E.

A thorough examination of the confusion/recognition tables also reveals a couple of trends. First, as expected, the Confusion/Recognition Model had either the best confusion/recognition values or values equally as good as the other models.

The next observation has the enumeration scheme giving a better confusion/recognition value than the Transmitted-Information Model in 24 cases, while the Transmitted-Information Model has better values in 22 cases. There were 73 instances where all three models gave the same optimal result (again, there were 149 total cases). So, in 73 cases, the Confusion/Recognition Model alone gave the optimal value.

Lastly, as with the information transmitted values, the amount of deviation in the results that were not equal did not appear to be significant from an absolute value standpoint in most cases. The importance of absolute optimality is determined by the application and the user of the data.

C. THE BOWEN DATA: A CLOSER LOOK

The Bowen data is of special interest because Bowen and his associates selected what they felt were the optimum subsets for subset sizes two through ten. Based on the article, their basis for selecting optimal subsets was confusion/recognition. Though these terms were not specifically used in this way, recognition was discussed, and the procedures used in the experiment did, in fact, base selection on the degree of recognition and confusion. For comparison purposes, the Bowen data is included in Table 52. (Bowen and others, 1960, pp. 28-30)

A quick scan of Table 52 reveals that Bowen and associates selected subsets very close in composition to those selected by the three models used in this thesis work. One of the most significant differences lies in their reluctance to use any of the symbols numbered higher than ten (except for symbol 14, the square). They didn't believe the higher numbered symbols were necessary because, as the number of the symbol increased, so did the degree of difficulty in recognizing the symbol. They did include the square in some of his optimal subsets, possibly due to a comfortable familiarity with the traditional, simple square. (Bowen and others, 1960, p.29)

The three models examined in this thesis produced results that were better, or as good as, the results of Bowen's experiment based on the indices used to evaluate optimality.

VII. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Before interpreting the results just discussed, it would be prudent to pause and examine the implications of drawing conclusions. Since human factors and human-system interface rely on human performance or human-system interaction, they are not precise sciences. Human interactions can be motivated by factors not easily integrated into formulas or models. Factors such as instinct, bias, and emotions are difficult, if not impossible, to predict. Some human reactions and interactions are fairly predictable, and as a result, human factors is a technical field of study. Still, the intangibles make dealing with some human factors issues difficult. However, the technology to bring optimal, or near optimal, solutions to problems such as these is available and provides a springboard for dealing with an inexact science.

What is optimal performance in the human factors environment? Or, what is the optimal solution to a problem dealing with human-system interface? As previously stated, the answers to these questions are best answered by the experts analyzing problems on a case by case basis. Fisher (in press) discusses two broad classes of optimization studies. In Type I studies, physical characteristics of design that affect optimal performance are the focus. In Type II studies, "...the goal is to identify the subset of design elements which optimize performance." The area of study covered by this thesis is Type II. He further discusses three classes used to organize the Type I and Type II studies: empirical, theoretical, and analytical. When there is a question concerning what optimality means or how it is to be used, Fisher's characterizations of optimization studies may provide an answer.

In this work, the objective was to develop a tool that a designer could use in system or concept design. The models developed simplify and standardize the selection of subsets that are optimal with respect to a given objective and given confusion matrix data. This brings up another potential problem area--the question of validity. Certainly, there is a desire to know if the models are valid. Sanders and McCormick (1987) discuss several types of validity: *face, content, and construct*. Face validity is concerned with whether a model appears to do what it was intended to do. Content validity pertains to whether the domain of interest is adequately represented or sampled. Construct validity asks whether the underlying essence of the actual problem is being addressed. They also discuss the concept of contamination in the measurement. Attention to these concepts early in the modeling process will help answer some of the questions that commonly arise such as: Was the data collection method sound? Was the experiment free from bias and noise? Were the test subjects qualified to perform as test subjects? Were they a properly diverse or properly restricted group (depending on the requirements)? Were they representative of the group affected by the outcome of the experiment?

These are important questions that can not be answered by examining the data sets. The experiment must be carefully controlled throughout. The models can only produce solutions based on the data given. The models can not anticipate, nor can they make judgements concerning the validity of the data.

The motivation behind this disclaimer is to ensure that more is not made of the models' capabilities than is warranted. The models will merely give a mathematically optimal--or near optimal, as the case may be--solution to the problem data given. With these ideas in mind, conclusions about the models' performance will be presented.

1. The Transmitted-Information Model

The Transmitted-Information Model developed in this thesis performed fairly well, but it did not consistently produce better results than the Confusion/Recognition Model. For information transmitted, the Confusion/Recognition Model actually performed better. As mentioned in the previous chapter, the Transmitted-Information Model returned a higher value of information transmitted than the Confusion/Recognition Model in 25 of 149 cases, while the

Confusion/Recognition Model returned a higher value in 30 cases. So, for these data sets, the Confusion/Recognition Model does a better job of maximizing information transmitted than the Transmitted-Information Model even though this is not the objective of the Confusion/Recognition Model. This is due to the unfortunate fact that the true information theory equations could not be fully implemented in the model because of their inherent nonlinearity. Recall that, the equations were boiled down to a single term. Considering this, the model performed quite well.

An interesting development was the performance of the program written in Turbo Pascal: the enumeration scheme. This model was intended as a check for the Transmitted-Information Model and was expected to return strictly better solutions since the Transmitted-Information Model was an approximation. But, it was anticipated that this program would use an inordinate amount of CPU time making it impractical for routine use. This was not the case.

The enumeration scheme solved the 15 smaller matrices to optimality in less than a minute. The Bowen data required approximately 24 hours to solve all possible subset sizes on an Intel 80386-based machine running at 33 MHz equipped with math coprocessor. Unfortunately, the attempt to solve the Moore data set was terminated after 24 hours of processing when it became evident that seven to ten days was going to be required for a complete solution.

A later attempt to process the Moore data set on an Intel 80486-based machine running at 33MHz proved more successful. The optimal solution for all subset sizes was completed in less than 48 hours. Solution times will probably improve dramatically within the next few years as technology pushes the speed of personal computers higher and higher. Another avenue of approach is processing on massively parallel computers capable of simultaneous processing on as many as 64,000 processors. This would be a very logical strategy for sets larger than the Moore set.

The solution times for the Transmitted-Information Model, using the previously mentioned 80386-based PC and GAMS version 2.25 with the XA solver, were very reasonable; no subset size for any of the data sets took more than about 15 minutes to solve. The longest solution times occurred for the Moore data set at subsets of size 11 through 14. The smaller data sets took on the order of one minute to provide solutions for all possible subset sizes.

Another interesting discovery was made in a review of the tables and is immediately obvious when viewing the graphs. In several data sets, as the subset size increased, the information transmitted began to decrease at some point. This can be interpreted as a decrease in system efficiency, or some may view it as information overload. Examining the confusion/recognition values will not reveal this system degradation in the way the Transmitted-Information Model or enumeration scheme do.

2. The Confusion/Recognition Model

The Confusion/Recognition Model outperformed the Transmitted-Information Model for both maximal information transmitted and minimal confusion with maximum recognition. However, the enumeration scheme outperformed the Confusion/Recognition Model for maximal information transmitted and did provide an insight into the previously mentioned reduction in efficiency. The solution times for the Confusion/Recognition Model were very reasonable, being about the same as those mentioned above for the Transmitted-Information Model.

B. RECOMMENDATIONS

Which model is best? It would be very nice to give a simple answer to this question, but this is not possible. One factor that influences the model of choice is the desires of the model user. Some may feel more comfortable with the information theory approach, while some may prefer the more intuitive confusion/recognition approach.

This brings up a point made by Wickens in his 1984 text. He lauds information theory as being a wide ranging theory "applicable across a wide variety of different dependent variables." (Wickens, 1984, pp.65-66) He later mentions criticisms of this theory including "limitations in the sensitivity of the information measure and limitations in its application to human performance." (Wickens, 1984, p.66) The second criticism dealing with applicability to human performance was discussed previously. The first criticism deals with the difference between consistency and correctness. Information theory will produce the same transmitted-information value for a situation where there is perfect recognition and where there is perfect confusion. As he points out, information theory must be used the with full awareness of the user. If the user does not check a model's solution, a "perfectly bad" subset may be used with the perception that it is "perfectly good". (Wickens, 1984, p.66)

If the information theory approach is chosen, the enumeration scheme should be used if possible since it provides optimal solutions with respect to maximal transmitted-information in all cases. If the data set is too large for the enumeration scheme and information theory is the desired approach, the Transmitted-Information Model may provide adequate results, although it will give sub-optimal results in many cases. The Transmitted-Information Model is not highly recommended.

Instead of the Transmitted-Information Model for larger data sets, the Confusion/Recognition Model is recommended. It bases optimality on an objective other than information transmitted but has been seen to provide better results with respect to information than the Transmitted-Information Model.

If the user wants to see any possible reductions in efficiency or information overloads, the Confusion/Recognition Model can produce the equivalent information transmitted value as a post-solve calculation. This data will reveal the desired insight as it did in this thesis. The Confusion/Recognition Model also bases optimality on a more easily grasped concept. For the average user, confusion and recognition may be more intuitive concepts. Also, recall that the time required for the enumeration scheme to run large data sets will become more tolerable as technology increases the speed of personal computers.

One of the goals of this thesis was also to determine if information theory and confusion theory would select the same optimal subsets. They didn't. The selected subsets were not different by a large degree. For this reason, the confusion/recognition values returned by the three models were not markedly different, nor were the transmitted-information values returned by the three model markedly different. In closing, either the Confusion/Recognition Model or the enumeration scheme will produce optimal results that are usable for most practical applications.

APPENDIX A INFORMATION THEORY MODEL (GAMS)

GAMS model for maximizing transmitted-information is presented here in edited form without comments or post-solve calculations so the entire model can be viewed at once. The full model used to generate the data in this thesis follows immediately afterward.

```
$TITLE INFORMATION THEORY MODEL
SETS  I   stimuli ;
ALIAS(I,J);
SCALAR  S  size of the subset to be selected ;
$INCLUDE SHEEHAN.DAT
SCALAR  T   total number of responses in matrix;
  T = SUM((I,J), C(I,J));
PARAMETER  P(I,J)
  P(I,J)  =  ( CARD(I) * C(I,J) )  /  (S* T) ;
PARAMETER  LP(I,J)  logarithmic probability matrix;
  LP(I,J) $ P(I,J)  =  P(I,J) * (LOG(1/P(I,J))/LOG(2));
BINARY VARIABLE
   X(I)      selected stimuli in subset ;
POSITIVE VARIABLE
   Y(I,J)    Indicator for joint selection of stimuli
FREE VARIABLE
   INFO       objective function value ;
EQUATIONS
   OBJFUNC           define objective function
   SUBSET            ensure proper subset size
   YDEF(I,J)         set y to one if i and j selected ;
SUBSET..    SUM(I, X(I)) =E= S ;
YDEF(I,J) $ (ord(i) lt ord(j))..  X(I) + X(J) - Y(I,J) =L= 1;
OBJFUNC..  SUM(I, LP(I,I) * X(I) )
           + SUM((I,J) $( ord(i) lt ord(j) ),
             Y(I,J) * ( LP(I,J) + LP(J,I) ) )
           =E= INFO ;
MODEL INFORM /ALL/;
LOOP(L,
 SOLVE INFORM USING MIP MINIMIZING INFO ;
 DISPLAY X.L ;
 S = S + 1;
 LNOW(L) = NO;
 LNOW(L + 1) = YES );
```

The complete model follows:

```
$TITLE INFORMATION THEORY MODEL
$offupper offsymxref offsymlist

* By Mike Sheehan 11/91  (Revised: RER 13 Nov 91)
* 2nd revision Mike Sheehan 12/91

OPTIONS
       limrow = 0
       limcol = 0
       solprint = off
       optcr =  0.0
       optca =  0.0
       iterlim = 100000
       reslim =  100000
       integer2 = 122
       integer1 = 1     ;

SETS  I   stimuli ;

ALIAS(I,J);

SCALAR  S  size of the subset to be selected ;

$INCLUDE SHEEHAN.DAT

SCALAR  T   total number of responses in matrix;

  T = SUM((I,J), C(I,J));


PARAMETER  P(I,J)  matrix of probabilities of each ;
*confusion value

  P(I,J)  =  ( CARD(I) * C(I,J) )  /  (S* T) ;


PARAMETER  LP(I,J)  logarithmic probability matrix;

  LP(I,J) $ P(I,J)  =  P(I,J) * (LOG(1/P(I,J))/LOG(2));


BINARY VARIABLE
   X(I)      selected stimuli in subset ;

POSITIVE VARIABLE
   Y(I,J)    Indicator for joint selection of stimuli
* y(i,j) is 1 if both x(i) and x(j) are 1 else y(i,j) is 0 ;
```

```
FREE VARIABLE
   INFO       objective function value ;

EQUATIONS

   OBJFUNC          define objective function
   SUBSET           ensure proper subset size
   YDEF(I,J)        set y to one if i and j selected ;

SUBSET..    SUM(I, X(I)) =E= S ;

YDEF(I,J) $ (ord(i) lt ord(j)).. X(I) + X(J) - Y(I,J) =L= 1;
*where i is less than j ensure y(i,j) is 1 only if both x(i)
*and x(j) are 1, for i greater than j is redundant


OBJFUNC..  SUM(I, LP(I,I) * X(I) )
*sum values of LP on main diagonal for chosen stimuli

         + SUM((I,J) $( ord(i) lt ord(j) ),
*sum values of LP where i is less than j and the i and j
*stimulus has been chosen

             Y(I,J) * ( LP(I,J) + LP(J,I) ) )
*sum values from LP matrix cells where i=j and j=i, this is
*equivalent to lower triangularizing the matrix (adding values
*from the i,j cell and j,i cell where i=j and j=i)

         =E= INFO ;


MODEL INFORM /ALL/;

PARAMETER

 CONFUSION(*,*) Confusion Among Selected Stimuli
 ENTROPY(*,*) Entropy Among Selected Stimuli
 NEWTOT total of all confusion values in selected
*       subset matrix;
 STIMPROB(I) probability of the i row in the
*       selected confusion matrix
 RESPPROB(J) probability of the j column in the
*       selected confusion matrix
 STIMINFO information derived from the stimuli
*       in the chosen subset
 RESPINFO information derived from the responses
*       in the chosen subset
 NEWLPMAT(I,J) logarithmic probability matrix using
*       values from chosen subset
 JOINTINFO joint information transmitted based on
*       chosen stimuli
```

```
 TOTALINFO total information transmitted based on
*       chosen stimuli (intersection of stim & resp info)
 RECOGNITN value of recognition for selected subset
*       based on Theise Mdl 3 included for comparison and
*       evaluation
;

LOOP(L,

SOLVE INFORM USING MIP MINIMIZING INFO ;

 CONFUSION(I,J) =  C(I,J) $( X.L(I) * X.L(J) ) ;

 ENTROPY(I,J) = LP(I,J) $( X.L(I) * X.L(J) ) ;

 NEWTOT = SUM((I,J), C(I,J)
                  $( X.L(I) * X.L(J) )) ;

 STIMPROB(I) = SUM(J, C(I,J)
                  $( X.L(I) * X.L(J) AND C(I,J) )/NEWTOT) ;

 RESPPROB(J) = SUM(I, C(I,J)
                  $( X.L(I) * X.L(J) AND C(I,J) )/NEWTOT) ;

 STIMINFO = SUM(I $ X.L(I),
            STIMPROB(I) * (LOG(1/STIMPROB(I))/LOG(2)));

 RESPINFO = SUM(J $ X.L(J),
            RESPPROB(J) * (LOG(1/RESPPROB(J))/LOG(2)));

 NEWLPMAT(I,J) $( X.L(I) * X.L(J) AND C(I,J) )
       = C(I,J)/NEWTOT * (( LOG(NEWTOT/C(I,J)))/LOG(2));

 JOINTINFO = SUM((I,J), NEWLPMAT(I,J) $( X.L(I)
                     * X.L(J) EQ 1 )) ;

 TOTALINFO = STIMINFO + RESPINFO - JOINTINFO;

 RECOGNITN = SUM(I $ X.L(I), C(I,I) );

DISPLAY X.L, RECOGNITN, TOTALINFO ;

S = S + 1;

LNOW(L) = NO;

LNOW(L + 1) = YES );
*end of loop
```

Sample input data file:

```
*WILPON9A.DAT - data file

SETS

      I   stimulus (rows)     /S0 * S9 /
      L   model runs            / RUN02 * RUN09 / ;

SCALAR  S  size of the subset to be selected  /2/ ;

TABLE C(I,*)      response j to stimulus i


       S0     S1     S2     S3     S4     S5     S6     S7     S8     S9
S0 63.8    0.0   12.5    0.0    5.7    0.0    3.6    6.8    0.0    0.0
S1  0.0   76.2    0.0    0.0   13.4    5.6    0.0    0.0    0.0    0.0
S2  0.0    0.0   66.8    5.4    0.0    0.0   12.7    4.2    8.0    0.0
S3  0.0    0.0    0.0   84.6    0.0    0.0    3.8    0.0    0.0    0.0
S4  5.0    0.0    3.4    0.0   88.5    0.0    0.0    0.0    0.0    0.0
S5  0.0    0.0    0.0    0.0    0.0   87.7    0.0    4.7    0.0    3.1
S6  0.0    0.0    0.0    5.8    0.0    0.0   72.1    3.5   15.5    0.0
S7  0.0    0.0    0.0    0.0    0.0    0.0    5.8   84.9    0.0    0.0
S8  0.0    0.0    0.0   10.0    0.0    0.0    7.9    5.6   72.5    0.0
S9  0.0    0.0    0.0    0.0    0.0   19.4    0.0   12.5    0.0   60.1
;
```

APPENDIX B CONFUSION MATRICES

Confusion matrices used as data sets:

CLARKE confusion matrix (Clarke, 1957, pp. 715-720)

	pa	ta	ka	fa	θa	sa
pa	405	242	162	128	048	015
ta	293	319	233	085	045	025
ka	208	440	240	023	057	032
fa	097	015	015	660	163	050
θa	058	050	040	315	340	197
sa	012	078	050	035	282	543

POLLACK1 confusion matrix (Pollack and Decker, 1960, pp.1-6)

	f	h	l	r	w	hw	y	#
f	96	0	0	1	2	0	0	0
h	6	84	0	0	0	0	0	9
l	1	1	76	12	5	2	2	0
r	1	1	11	57	14	5	11	0
w	1	0	3	5	69	15	8	0
hw	1	1	2	3	25	62	7	0
y	0	1	1	1	3	1	94	0
#	2	6	0	0	1	0	0	91

POLLACK2 confusion matrix

	f	h	l	r	w	hw	y	#
f	89	2	1	2	2	3	1	0
h	14	70	1	1	1	0	0	12
l	4	3	63	8	12	4	5	1
r	1	1	8	40	25	10	16	0
w	1	0	2	7	61	20	8	1
hw	5	1	1	1	20	65	8	0
y	1	1	6	7	12	2	71	0
#	3	8	0	0	0	0	1	88

POLLACK3 confusion matrix

	f	h	l	r	w	hw	y	#
f	66	10	4	4	4	4	2	5
h	14	54	4	2	2	2	1	21
l	4	3	48	12	16	7	6	3
r	3	3	20	27	25	9	11	1
w	4	2	10	13	48	12	11	0
hw	9	3	4	6	26	42	10	1
y	1	2	16	12	22	7	40	1
#	8	20	4	3	3	2	1	60

POLLACK4 confusion matrix

	f	h	l	r	w	hw	y	#
f	28	20	12	4	7	4	3	22
h	8	45	14	3	7	2	6	15
l	6	7	34	7	17	13	9	8
r	2	7	20	18	26	8	11	8
w	5	7	17	11	28	9	15	9
hw	9	8	13	9	17	27	9	7
y	3	6	17	14	23	12	19	6
#	13	30	9	3	4	3	6	32

WILPON10 confusion matrix (Wilpon, 1985, pp. 423-451)

	0	1	2	3	4	5	6	7	8	9
0	86.5	0.0	0.0	0.0	5.6	0.0	0.0	0.0	0.0	0.0
1	0.0	94.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	90.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	93.9	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	94.4	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	92.5	0.0	0.0	0.0	3.4
6	0.0	0.0	0.0	0.0	0.0	0.0	85.7	0.0	7.1	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	92.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	3.3	0.0	90.5	0.0
9	0.0	0.0	0.0	0.0	0.0	7.5	0.0	0.0	0.0	84.2

WILPON7A confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	69.6	0.0	0.0	0.0	15.6	0.0	0.0	5.1	0.0	0.0
1	0.0	88.2	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0
2	4.6	0.0	78.2	0.0	0.0	0.0	4.7	5.4	0.0	0.0
3	0.0	0.0	0.0	91.3	0.0	0.0	3.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	95.4	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	87.8	0.0	0.0	0.0	6.9
6	0.0	0.0	0.0	4.2	0.0	0.0	79.3	0.0	11.6	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	88.4	0.0	0.0
8	0.0	0.0	0.0	7.5	0.0	0.0	5.6	0.0	81.2	0.0
9	0.0	3.1	0.0	0.0	0.0	12.8	0.0	4.6	0.0	74.9

WILPON7B confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	66.3	0.0	0.0	0.0	27.7	0.0	0.0	0.0	0.0	0.0
1	0.0	94.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	8.1	0.0	77.8	0.0	8.1	0.0	0.0	4.5	0.0	0.0
3	0.0	0.0	0.0	95.7	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	3.3	0.0	0.0	93.6	0.0	0.0	0.0	0.0	0.0
5	0.0	6.8	0.0	0.0	0.0	84.0	0.0	0.0	0.0	5.0
6	0.0	0.0	0.0	0.0	0.0	0.0	82.4	6.8	5.1	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	85.5	0.0	5.0
8	0.0	0.0	0.0	0.0	0.0	0.0	5.8	0.0	90.3	0.0
9	0.0	4.4	0.0	4.1	0.0	8.9	0.0	0.0	0.0	79.0

WILPON7C confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	98.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
5	0.0	4.0	0.0	0.0	5.0	75.0	0.0	4.0	0.0	11.0
6	0.0	0.0	3.0	0.0	0.0	0.0	94.0	0.0	0.0	0.0
7	0.0	0.0	10.0	0.0	0.0	0.0	0.0	87.0	0.0	5.0
8	3.0	0.0	3.0	3.0	0.0	0.0	4.0	0.0	87.0	0.0
9	0.0	9.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	84.0

WILPON8A confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	58.4	0.0	11.8	0.0	0.0	0.0	4.7	11.8	6.5	0.0
1	7.8	46.3	0.0	0.0	6.4	20.1	0.0	4.0	0.0	8.3
2	0.0	0.0	47.9	3.3	0.0	0.0	7.0	19.4	19.9	0.0
3	0.0	0.0	0.0	74.2	0.0	0.0	7.0	5.4	7.5	0.0
4	28.6	3.1	0.0	0.0	50.6	8.5	0.0	3.8	0.0	0.0
5	3.3	0.0	0.0	0.0	0.0	79.6	7.4	4.2	0.0	3.9
6	0.0	0.0	0.0	4.0	0.0	0.0	62.6	5.0	24.9	0.0
7	0.0	0.0	4.6	0.0	0.0	3.0	12.3	69.4	5.1	3.0
8	0.0	0.0	0.0	7.4	0.0	0.0	0.0	4.7	79.2	0.0
9	0.0	0.0	0.0	0.0	0.0	26.7	14.7	10.2	0.0	43.2

WILPON8B confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	84.3	0.0	5.6	0.0	0.0	0.0	0.0	3.3	0.0	0.0
1	6.3	72.9	0.0	0.0	0.0	7.8	0.0	0.0	4.3	6.3
2	0.0	0.0	86.4	0.0	0.0	0.0	0.0	6.0	0.0	0.0
3	0.0	0.0	0.0	87.7	0.0	0.0	0.0	0.0	5.4	0.0
4	34.0	0.0	0.0	0.0	48.7	8.8	0.0	0.0	0.0	0.0
5	3.2	0.0	0.0	0.0	0.0	80.4	5.6	0.0	0.0	6.5
6	0.0	0.0	0.0	0.0	0.0	0.0	85.8	0.0	7.5	0.0
7	0.0	0.0	0.0	0.0	0.0	3.2	9.2	74.8	3.0	3.2
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	95.3	0.0
9	0.0	0.0	0.0	4.4	0.0	12.6	12.3	3.9	0.0	64.3

WILPON8C confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	99.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	97.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	0.0	91.0	6.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.0

WILPON9A confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	63.8	0.0	12.5	0.0	5.7	0.0	3.6	6.8	0.0	0.0
1	0.0	76.2	0.0	0.0	13.4	5.6	0.0	0.0	0.0	0.0
2	0.0	0.0	66.8	5.4	0.0	0.0	12.7	4.2	8.0	0.0
3	0.0	0.0	0.0	84.6	0.0	0.0	3.8	0.0	0.0	0.0
4	5.0	0.0	3.4	0.0	88.5	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	87.7	0.0	4.7	0.0	3.1
6	0.0	0.0	0.0	5.8	0.0	0.0	72.1	3.5	15.5	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	5.8	84.9	0.0	0.0
8	0.0	0.0	0.0	10.0	0.0	0.0	7.9	5.6	72.5	0.0
9	0.0	0.0	0.0	0.0	0.0	19.4	0.0	12.5	0.0	60.1

WILPON9B confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	90.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	95.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	95.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	94.2	0.0	0.0	0.0	0.0	0.0	0.0
4	3.8	0.0	0.0	0.0	93.4	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	93.1	0.0	0.0	0.0	3.0
6	0.0	0.0	0.0	0.0	0.0	0.0	87.0	3.2	4.2	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	93.3	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	95.3	0.0
9	0.0	0.0	0.0	0.0	0.0	9.2	0.0	0.0	0.0	85.0

WILPON9C confusion matrix

	0	1	2	3	4	5	6	7	8	9
0	87.0	0.0	9.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0
1	0.0	98.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	98.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	98.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	3.0	0.0	97.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	8.0	72.0	0.0	3.0	0.0	14.0
6	0.0	0.0	4.0	0.0	0.0	0.0	91.0	5.0	0.0	0.0
7	0.0	0.0	7.0	0.0	0.0	0.0	0.0	91.0	0.0	0.0
8	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	94.0	0.0
9	0.0	7.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90.0

BOWEN confusion matrix

0.916	0.000	0.000	0.000	0.000	0.006	0.000	0.012	0.000	0.000	0.006	0.012	0.006	0.351	0.006	0.012	0.048	0.024	0.006	0.024
0.000	0.898	0.000	0.000	0.012	0.006	0.006	0.012	0.006	0.006	0.006	0.024	0.000	0.018	0.006	0.006	0.012	0.060	0.000	0.030
0.000	0.000	0.869	0.006	0.000	0.006	0.024	0.000	0.006	0.000	0.060	0.000	0.012	0.006	0.036	0.024	0.125	0.006	0.030	0.000
0.006	0.000	0.006	0.869	0.000	0.000	0.000	0.000	0.000	0.042	0.000	0.006	0.006	0.018	0.042	0.000	0.006	0.012	0.000	0.000
0.006	0.000	0.000	0.012	0.839	0.006	0.000	0.018	0.006	0.012	0.006	0.048	0.000	0.000	0.000	0.006	0.000	0.024	0.000	0.137
0.006	0.006	0.006	0.000	0.000	0.881	0.000	0.000	0.012	0.012	0.006	0.006	0.000	0.012	0.024	0.000	0.012	0.000	0.012	0.000
0.000	0.012	0.000	0.006	0.000	0.000	0.875	0.000	0.000	0.000	0.006	0.048	0.024	0.000	0.018	0.149	0.006	0.006	0.208	0.006
0.006	0.012	0.012	0.000	0.012	0.006	0.000	0.833	0.006	0.006	0.000	0.012	0.048	0.012	0.006	0.018	0.006	0.095	0.000	0.012
0.006	0.006	0.012	0.030	0.012	0.012	0.024	0.000	0.839	0.012	0.030	0.018	0.024	0.000	0.024	0.018	0.030	0.006	0.024	0.006
0.012	0.000	0.018	0.000	0.000	0.018	0.018	0.006	0.024	0.863	0.000	0.006	0.006	0.006	0.006	0.000	0.012	0.000	0.012	0.006
0.000	0.000	0.036	0.018	0.012	0.012	0.006	0.012	0.012	0.006	0.785	0.000	0.000	0.006	0.048	0.000	0.036	0.000	0.018	0.012
0.006	0.012	0.000	0.006	0.024	0.000	0.018	0.030	0.006	0.006	0.012	0.756	0.000	0.006	0.006	0.006	0.060	0.018	0.012	0.000
0.000	0.006	0.006	0.000	0.012	0.012	0.000	0.000	0.000	0.006	0.006	0.006	0.779	0.012	0.000	0.006	0.012	0.006	0.006	0.000
0.000	0.006	0.000	0.000	0.000	0.006	0.000	0.012	0.000	0.000	0.006	0.000	0.012	0.506	0.000	0.000	0.113	0.006	0.006	0.018
0.012	0.000	0.018	0.012	0.012	0.012	0.000	0.006	0.012	0.012	0.036	0.000	0.006	0.006	0.762	0.000	0.018	0.000	0.018	0.006
0.000	0.000	0.000	0.006	0.000	0.000	0.006	0.018	0.000	0.000	0.000	0.006	0.000	0.000	0.006	0.553	0.006	0.000	0.065	0.006
0.018	0.006	0.006	0.000	0.000	0.012	0.006	0.012	0.042	0.012	0.006	0.012	0.018	0.006	0.006	0.030	0.458	0.000	0.018	0.006
0.000	0.036	0.000	0.006	0.000	0.000	0.006	0.024	0.006	0.006	0.000	0.030	0.030	0.012	0.000	0.006	0.018	0.720	0.000	0.012
0.000	0.000	0.012	0.018	0.006	0.000	0.006	0.000	0.018	0.000	0.018	0.006	0.018	0.006	0.006	0.167	0.018	0.006	0.559	0.024
0.006	0.000	0.000	0.012	0.060	0.006	0.006	0.006	0.006	0.000	0.012	0.006	0.006	0.018	0.000	0.000	0.006	0.012	0.000	0.690

APPENDIX C ENUMERATION SCHEME (TURBO PASCAL)

Listing for Turbo Pascal program called INFO:

```
program information(infile,outfile);

type
   rangearray = array[1..35] of integer;
   sqarray = array[1..35, 1..35] of real;
   stname = string[5];
   smallsub = array[1..5] of real;
var
   i, j, k, subsetsize, stim : integer;
   ln2, count : real;
   subset : rangearray;
   confusion : sqarray;
   infoin : string[8];
   infile, outfile : text;
   stimname : array[1..35] of stname;
   subsetname : array[1..35] of stname;
   topfive : smallsub;
   tfsubset : array[1..35, 1..5]of stname;

function totalinfo(subset:rangearray) : real;
   var rowinfo, colinfo, jointinfo : real;
   var rowtot, coltot, mattotal, jointprob : real;

   begin
   mattotal := 0;
   for i := 1 to subsetsize do
      begin
      for j := 1 to subsetsize do
         mattotal := mattotal +
                  confusion[subset[i],subset[j]];
      end;
   jointinfo := 0;
   rowinfo := 0;
   colinfo := 0;
   for i := 1 to subsetsize do
      begin
      rowtot := 0;
      coltot := 0;
      for j := 1 to subsetsize do
         begin
         rowtot := rowtot + confusion[subset[i],subset[j]];
```

```
          coltot := coltot + confusion[subset[j],subset[i]];
          jointprob := confusion[subset[i],subset[j]]/mattotal;
          if jointprob <> 0 then
          jointinfo := jointinfo - (jointprob) *
                                            (ln(jointprob)/ln2);
          end;
       rowinfo := rowinfo - rowtot/mattotal *
                                      (ln(rowtot/mattotal)/ln2);
       colinfo := colinfo - coltot/mattotal *
                                      (ln(coltot/mattotal)/ln2);
       end;
    totalinfo := rowinfo + colinfo - jointinfo;
    end  { function "totalinfo" };


 procedure evaluate(var val : real);
    var i ,j, k : integer;
    var temp : real;
    var tempset : array[1..35] of stname;

    begin
    for i := 1 to 5 do
       begin
       if val > topfive[i] then
          begin
          temp := topfive[i];
          for k := 1 to 35 do
             tempset[k] := tfsubset[k,i];
          topfive[i] := val;
          for k := 1 to 35 do
             tfsubset[k,i] := subsetname[k];
          val := temp;
          for k := 1 to 35 do
             subsetname[k] := tempset[k];
          end  { if loop };
       end  { for loop };
    end  {procedure "evaluate" };


 procedure process(subset:rangearray; size:integer);
    var j:integer;
    var value : real;

    begin
    count := count + 1;
    for j:= 1 to subsetsize do
       subsetname[j] := stimname[subset[j]];
    value := totalinfo(subset);
    evaluate(value);
    end { procedure "process" };
```

```
procedure combs(n,r:integer)  {(Rohl, 1983, pp. 154-157)};
   var s: rangearray;

   procedure choose(d,lower:integer);
      var i:integer;
      begin
      for i:= lower to n-r+d do
         begin
         s[d] := i;
         if d <> r then choose(d+1,i+1) else process(s,r)
      end { of loop on "i" }
   end { of procedure "choose" };

begin
choose(1,1)
end { of procedure "combs" };


procedure storeinfo(size: integer);
var i, j : integer;

begin  { procedure storeinfo }
write(outfile, 'The number of subsets of size ', size, '
                                    examined was: ');
writeln(outfile, count:8:0);
write(outfile, 'The following 5 subsets had the highest
                                         info');
writeln(outfile,' transfer values.');
for i := 1 to 5 do
   begin
   for j := 1 to size do
      write(outfile, tfsubset[j,i], ' ');
   writeln(outfile);
   writeln(outfile, 'Info transmitted:  ', topfive[i] :7:4);
   writeln(outfile);
   end  { for loop };
end  { procedure storeinfo };


procedure getdata(var stimuli :integer);
var i, j, no_lines : integer;

begin  { procedure getdata }
reset(infile);
no_lines := 0;
while not EOF(infile) do
   begin
   no_lines := no_lines + 1;
   readln(infile);
   end;
stimuli := no_lines div 2;
```

```
reset(infile);
for i := 1 to stimuli do
   readln(infile,stimname[i]);
for i := 1 to stimuli do
   begin
   writeln;
   for j := 1 to stimuli do
      begin
      read(infile, confusion[i,j]);
      write(confusion[i,j]:5:2,' ');
      end;
   readln(infile);
   end  { for loop };
writeln;
end  { procedure getdata };


begin { MAIN PROGRAM }
ln2 := ln(2);
write('What file do you want to process (8 character name)?');
readln(infoin);
assign(infile, concat(infoin, '.dat'));
assign(outfile, concat(infoin, '.out'));
rewrite(outfile);
writeln(outfile,'This data file is called:
                                 ',infoin + '.DAT');
writeln;
writeln('This data file is called:  ',infoin + '.DAT');
getdata(stim);
for i:= 2 to stim do
   begin
   count := 0;
   subsetsize := i;
   for j := 1 to 5 do
      begin
      topfive[j] := 0;
      for k := 1 to 35 do
         tfsubset[k,j] := '0';
      end  { for loop };
   writeln;
   writeln('Now processing subsets of size ',i);
   combs(stim,i);
   writeln('Done with subsets of size ',i);
   writeln('There were ',count:8:0,' subsets of size ',i);
   storeinfo(i);
   end;
close(infile);
close(outfile);
end.
```

Sample input data file (WILPON9A):

```
0
1
2
3
4
5
6
7
8
9
63.8  0.0   12.5  0.0   5.7   0.0   3.6   6.8   0.0   0.0   0.0
0.0   76.2  0.0   0.0   13.4  5.6   0.0   0.0   0.0   0.0   0.0
0.0   0.0   66.8  5.4   0.0   0.0   12.7  4.2   8.0   0.0   0.0
0.0   0.0   0.0   84.6  0.0   0.0   3.8   0.0   0.0   0.0   0.0
5.0   0.0   3.4   0.0   88.5  0.0   0.0   0.0   0.0   0.0   0.0
0.0   0.0   0.0   0.0   0.0   87.7  0.0   4.7   0.0   3.1   0.0
0.0   0.0   0.0   5.8   0.0   0.0   72.1  3.5   15.5  0.0   0.0
0.0   0.0   0.0   0.0   0.0   0.0   5.8   84.9  0.0   0.0   0.0
0.0   0.0   0.0   10.0  0.0   0.0   7.9   5.6   72.5  0.0   0.0
0.0   0.0   0.0   0.0   0.0   19.4  0.0   12.5  0.0   60.1  0.0
```

APPENDIX D THE CONFUSION/RECOGNITION MODEL (MODIFIED)

Theise's GAMS model for maximizing recognition while minimizing confusion:

```
$TITLE THEISE RECOGNITION MODEL -
$offupper offsymxref offsymlist

* Revision By Mike Sheehan 2/92

OPTIONS
       limrow = 0
       limcol = 0
       solprint = off
       optcr =  0.0
       optca =  0.00
       iterlim = 100000
       reslim =  100000
       integer2 = 122
       integer1 = 1    ;

SETS  I   stimuli ;

ALIAS(I,J);

SCALAR  S  size of the subset to be selected ;

$INCLUDE THEISE.DAT

SCALAR  M   total number of responses in matrix ;

  M = SUM((I,J), C(I,J));

PARAMETER  P(I,J)  matrix of prob of each confusion value;

  P(I,J)  =  ( CARD(I) * C(I,J) )  /  (S* M) ;

PARAMETER  LP(I,J)  logarithmic probability matrix;

  LP(I,J) $ P(I,J)  =  P(I,J) * (LOG(1/P(I,J))/LOG(2));

BINARY VARIABLE
   X(I)      selected stimuli in subset  ;
```

```
POSITIVE VARIABLE
      Y(I,J)     Indicator for joint selection of stimuli
* y(i,j) is 1 if both x(i) and x(j) are 1 else y(i,j) is 0;

FREE VARIABLE
   DPLUS   deviation from confusion threshold
   REC       objective function value    ;

EQUATIONS

   OBJFUNC          define objective function
   SUBSET           ensure proper subset size
   YDEF(I,J)        set y to one if i and j selected
   CONFUSE          ensure minimum confusion   ;

SUBSET..     SUM(I, X(I)) =E= S ;

YDEF(I,J)$(ord(i) lt ord(j))..  X(I) + X(J) - Y(I,J) =L= 1 ;
*where i is less than j ensure y(i,j) is 1 iff both x(i) and
*x(j) are 1, for i greater than j is redundant

CONFUSE..     SUM((I,J) $ (ORD(I) LT ORD(J)),
*sum values of confusion in upper triangle of matrix

             (C(I,J) + C(J,I)) * Y(I,J))
*add values of confusion from complementary cells in matrix
*effectively upper triangularizes the matrix

             + SUM(I, U(I) * X(I)) - DPLUS =L= T ;
*add relavent values of u (non-responses) then ensure the
*confusion value is less than (or equal to) threshold value
*if not, variable dplus will conpensate for the difference
*and ensure the inequality condition holds

OBJFUNC..  REC =E= SUM(I, C(I,I) * X(I) - M * DPLUS ) ;
*sum values of C on main diagonal for chosen stimuli
*then subtract deviation from confusion threshold times
*large constant

MODEL RECOG /ALL/;

PARAMETER ENTROPY(*,*) Entropy Among Selected Stimuli ;
PARAMETER NEWTOT total of confusion values in chosen matrix;
PARAMETER STIMPROB(I) probability of the i row in selected ;
*confusion matrix
PARAMETER RESPPROB(J) probability of the j column in the ;
*selected confusion matrix
PARAMETER STIMINFO information derived from the stimuli;
*in the chosen subet
PARAMETER RESPINFO information derived from the responses;
*in the chosen subset
```

```
PARAMETER NEWLPMAT(I,J) logarithmic probability matrix using;
*values from chosen subset
PARAMETER JOINTINFO joint information transmitted based on;
*chosen stimuli
PARAMETER TOTALINFO total information transmitted based on;
*chosen stimuli (intersection of stimulus & response info)
PARAMETER RECOGNITN value of recognition for selected stimuli;
PARAMETER CONFUSION post solve to calc confusion;

LOOP(L,

SOLVE RECOG USING MIP MAXIMIZING REC ;

 CONFUSION =  SUM((I,J) $ (ORD(I) lt ORD(J)),
                  (X.L(I) * X.L(J)) * ( C(I,J) + C(J,I) )) ;

 ENTROPY(I,J) = LP(I,J) $( X.L(I) * X.L(J) ) ;

 NEWTOT = SUM((I,J), C(I,J)
                    $( X.L(I) * X.L(J) )) ;

 STIMPROB(I) = SUM(J, C(I,J)
                    $( X.L(I) * X.L(J) AND C(I,J) )/NEWTOT) ;

 RESPPROB(J) = SUM(I, C(I,J)
                    $( X.L(I) * X.L(J) AND C(I,J) )/NEWTOT) ;

 STIMINFO = SUM(I $ X.L(I),
              STIMPROB(I) * (LOG(1/STIMPROB(I))/LOG(2)));

 RESPINFO = SUM(J $ X.L(J),
              RESPPROB(J) * (LOG(1/RESPPROB(J))/LOG(2)));

 NEWLPMAT(I,J) $( X.L(I) * X.L(J) AND C(I,J) )
        = C(I,J)/NEWTOT * (( LOG(NEWTOT/C(I,J)))/LOG(2));

 JOINTINFO = SUM((I,J), NEWLPMAT(I,J) $( X.L(I)
                        * X.L(J) EQ 1 )) ;

 TOTALINFO = STIMINFO + RESPINFO - JOINTINFO;

 RECOGNITN = SUM(I $ X.L(I), C(I,I) ) ;

DISPLAY X.L, DPLUS.L, M, TOTALINFO, CONFUSION, RECOGNITN ;

S = S + 1;

LNOW(L) = NO;

LNOW(L + 1) = YES );
*end of loop
```

Input data file for the Confusion/Recognition Model:

```
SETS

      I   stimulus (rows)      /S0 * S9 /
      L   model runs             / RUN02 * RUN10 / ;

SCALAR  S  size of the subset to be selected  /2/ ;

SET  LNOW(L)   dynamic set for current run  / RUN02 /;

SCALAR  T  confusion threshold       / 0 / ;

PARAMETER U(I)   nonresponses in confusion matrix
        / S0  0 /;

TABLE C(I,*)     response j to stimulus i

      S0    S1    S2    S3    S4    S5    S6    S7    S8    S9    U
S0  63.8   0.0  12.5   0.0   5.7   0.0   3.6   6.8   0.0   0.0   0.0
S1   0.0  76.2   0.0   0.0  13.4   5.6   0.0   0.0   0.0   0.0   0.0
S2   0.0   0.0  66.8   5.4   0.0   0.0  12.7   4.2   8.0   0.0   0.0
S3   0.0   0.0   0.0  84.6   0.0   0.0   3.8   0.0   0.0   0.0   0.0
S4   5.0   0.0   3.4   0.0  88.5   0.0   0.0   0.0   0.0   0.0   0.0
S5   0.0   0.0   0.0   0.0   0.0  87.7   0.0   4.7   0.0   3.1   0.0
S6   0.0   0.0   0.0   5.8   0.0   0.0  72.1   3.5  15.5   0.0   0.0
S7   0.0   0.0   0.0   0.0   0.0   0.0   5.8  84.9   0.0   0.0   0.0
S8   0.0   0.0   0.0  10.0   0.0   0.0   7.9   5.6  72.5   0.0   0.0
S9   0.0   0.0   0.0   0.0   0.0  19.4   0.0  12.5   0.0  60.1   0.0
;

*WILPON9A.DAT - data file
```

APPENDIX E DATA COMPARISON TABLES AND GRAPHS

Tables and graphs compiling output data from the three models:

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	ka, fa	38	900	0.613
3	ka, fa, sa	205	1443	0.856
4	pa, ka, fa, sa	827	1848	0.791
5	pa, ka, fa, Oa, sa	1987	2188	0.599
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	pa, sa	27	948	0.803
3	ka, fa, sa	205	1443	0.856
4	ta, ka, fa, sa	827	1848	0.791
5	pa, ta, ka, fa, sa	1987	2188	0.599
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	pa, sa	27	948	0.803
3	ka, fa, sa	205	1443	0.856
4	pa, ka, fa, sa	1081	1762	0.817
5	pa, ka, fa, Oa, sa	2238	2167	0.651

Figure 3 Comparison of Model Results for Clarke Data Set

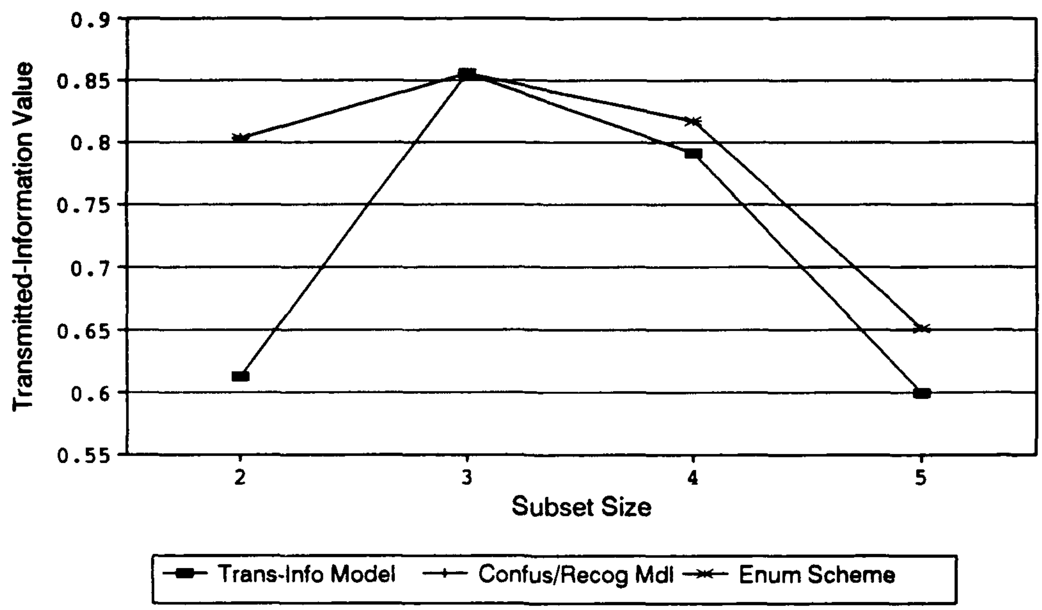

Figure 4 Clarke Data Set: Transmitted-Information

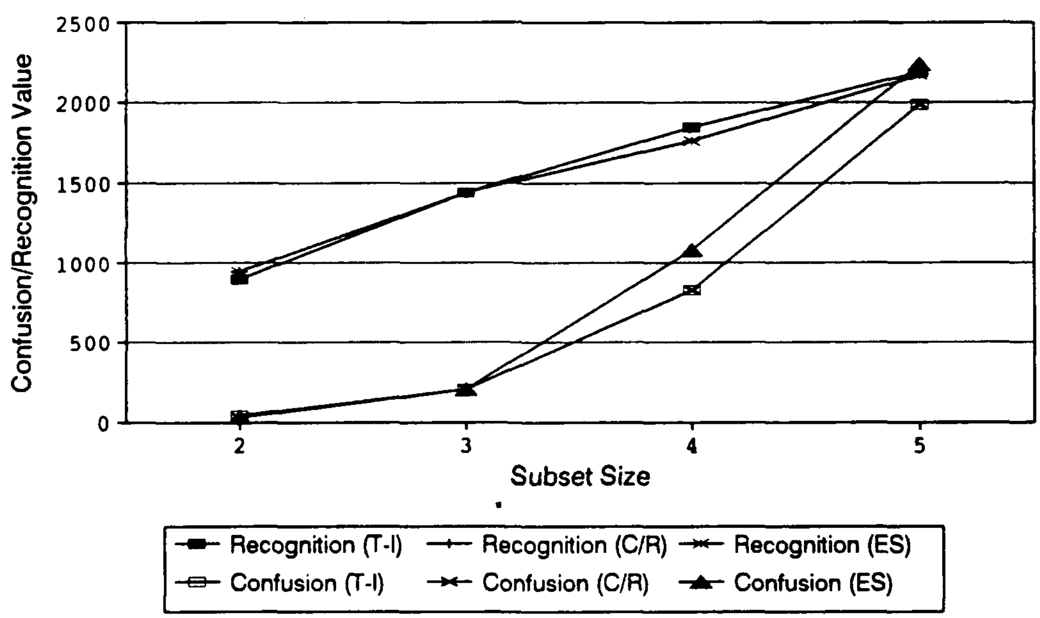

Figure 5 Clarke Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/	Recognition	Transmitted information
2	f, y	0	190	1.000
3	f, y, #	2	281	1.535
4	f, l, y, #	6	357	1.875
5	f, l, hw, y, #	19	419	2.033
6	f, h, l, hw, y, #	43	503	2.109
7	f, h, l, r, hw, y, #	89	560	2.067

Subset size	Confusion/Recognition Model Selected Subsets	Confusion/	Recognition	Transmitted information
2	f, y	0	190	1.000
3	f, y, #	2	281	1.535
4	f, l, y, #	6	357	1.875
5	f, l, hw, y, #	19	419	2.033
6	f, h, l, hw, y, #	43	503	2.109
7	f, h, l, r, hw, y, #	89	560	2.067

Subset size	Enumeration Scheme Selected Subsets	Confusion/	Recognition	Transmitted information
2	f, y	0	190	1.000
3	f, y, #	2	281	1.535
4	f, l, y, #	6	357	1.875
5	f, l, hw, y, #	19	419	2.033
6	f, h, l, hw, y, #	43	503	2.109
7	f, h, l, r, hw, y, #	89	560	2.067

Figure 6 Comparison of Model Results for Pollack1 Data Set

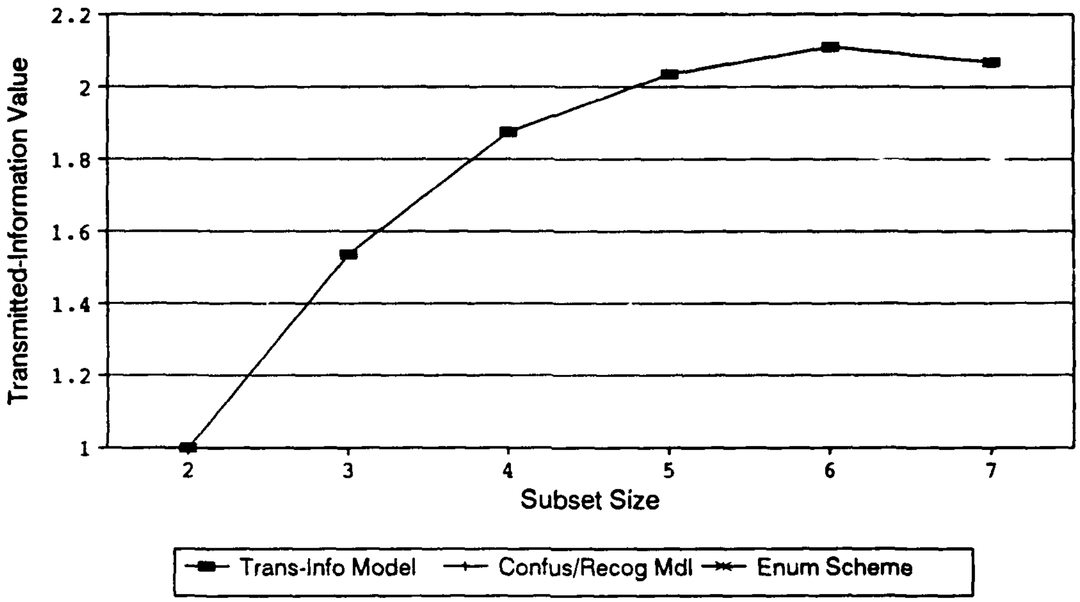

Figure 7 Pollack1 Data Set: Transmitted-Information

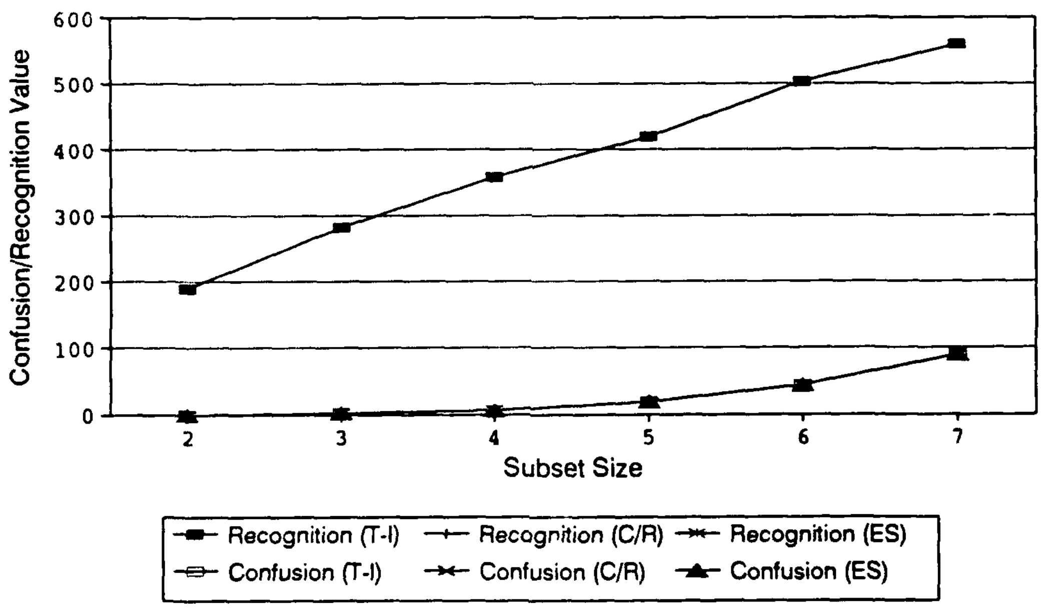

Figure 8 Pollack1 Data Set: Confusion/Recognition

Subset size	*Transmitted-Information Model Selected Subsets*	Confusion/ Recognition		*Transmitted information*
2	r, #	0	128	0.896
3	f, r, #	6	217	1.331
4	f, r, hw, #	25	282	1.522
5	f, l, hw, y, #	46	376	1.686
6	f, h, l, hw, y, #	88	446	1.692
7	f, h, l, r, hw, y, #	143	486	1.650
Subset size	*Confusion/Recognition Model Selected Subsets*	*Confusion/ Recognition*		*Transmitted information*
2	hw, #	0	153	0.984
3	f, y, #	6	248	1.410
4	f, l, hw, #	22	305	1.577
5	f, l, hw, y, #	46	376	1.686
6	f, h, l, hw, y, #	88	446	1.691
7	f, h, l, r, hw, y, #	143	486	1.650
Subset size	*Enumeration Scheme Selected Subsets*	*Confusion/ Recognition*		*Transmitted information*
2	hw, #	0	153	0.984
3	f, y, #	6	248	1.410
4	f, l, y, #	23	311	1.580
5	f, l, hw, y, #	46	376	1.686
6	f, h, l, hw, y, #	88	446	1.692
7	f, h, l, r, hw, y, #	143	486	1.650

Figure 9 Comparison of Model Results for Pollack2 Data Set

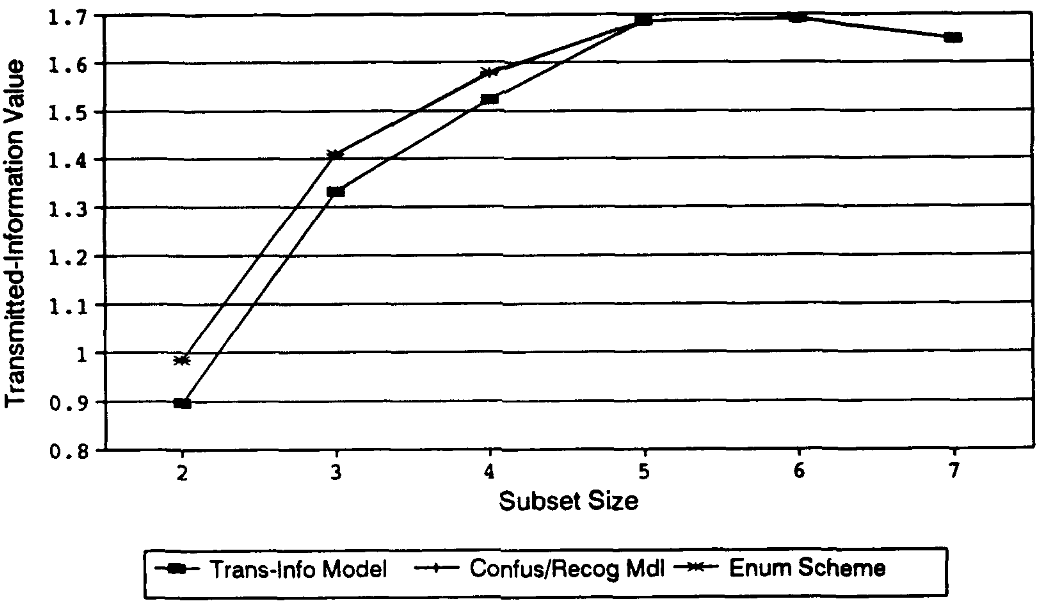

Figure 10 Pollack2 Data Set: Transmitted-Information

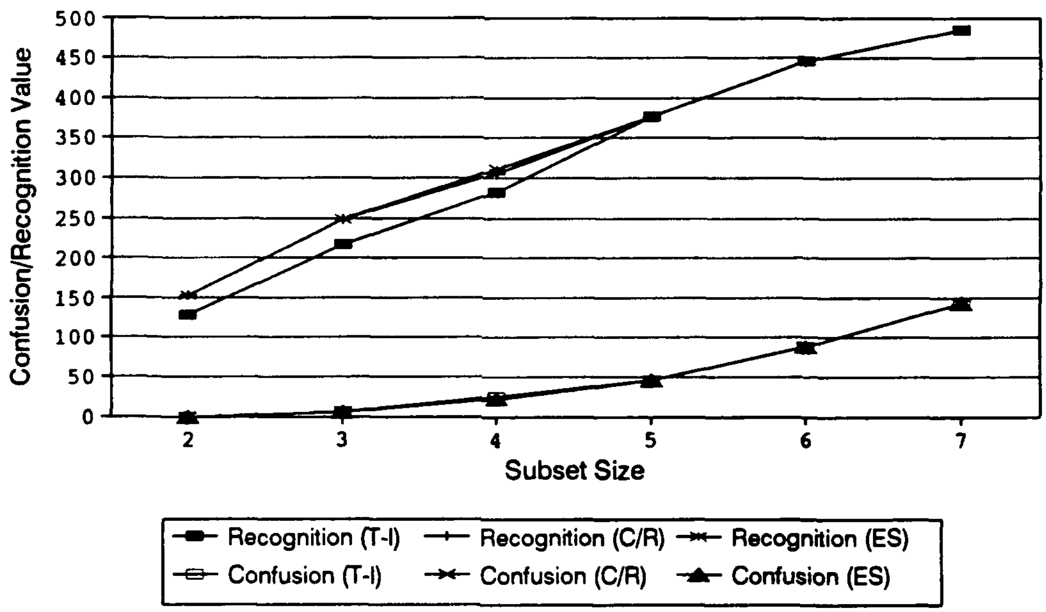

Figure 11 Pollack2 Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	r, #	4	87	0.655
3	r, hw, #	22	129	0.893
4	f, r, y, #	52	193	1.019
5	f, r, hw, y, #	100	235	0.972
6	f, h, r, hw, y, #	178	289	0.914
7	f, h, l, r, hw, y, #	265	337	0.886
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	y, #	2	100	0.833
3	f, y, #	18	166	1.027
4	f, hw, y, #	51	208	1.041
5	f, l, hw, y, #	99	256	1.013
6	f, h, r, hw, y, #	178	289	0.914
7	f, h, l, r, hw, y, #	265	337	0.886
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	w, #	3	108	0.839
3	f, y, #	18	166	1.027
4	f, w, y, #	62	214	1.048
5	f, l, hw, y, #	99	256	1.013
6	f, h, l, hw, y, #	179	310	0.944
7	f, h, l, w, hw, y, #	291	358	0.895

Figure 12 Comparison of Model Results for Pollack3 Data Set

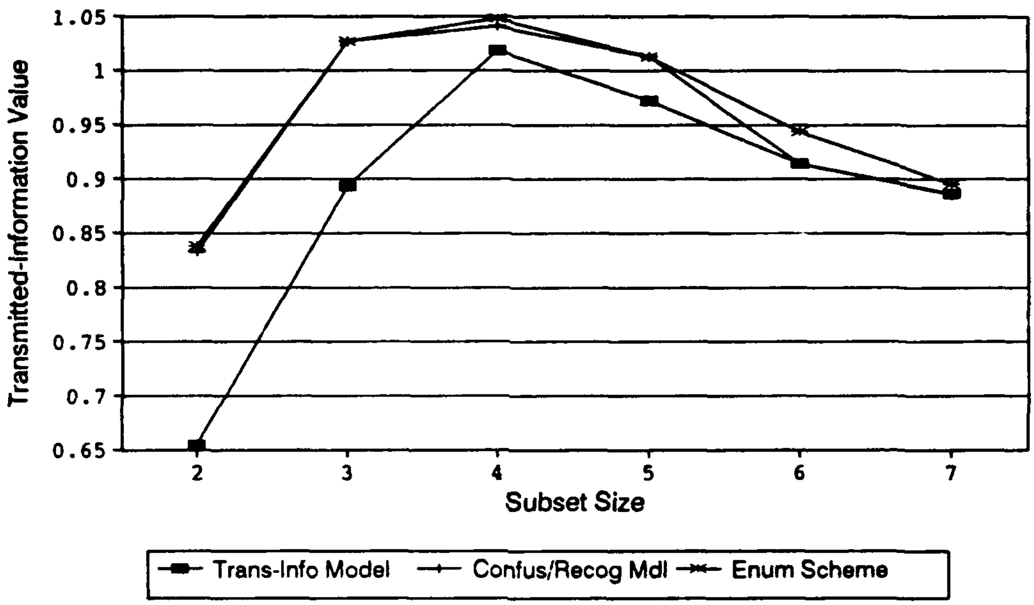

Figure 13 Pollack3 Data Set: Transmitted-Information

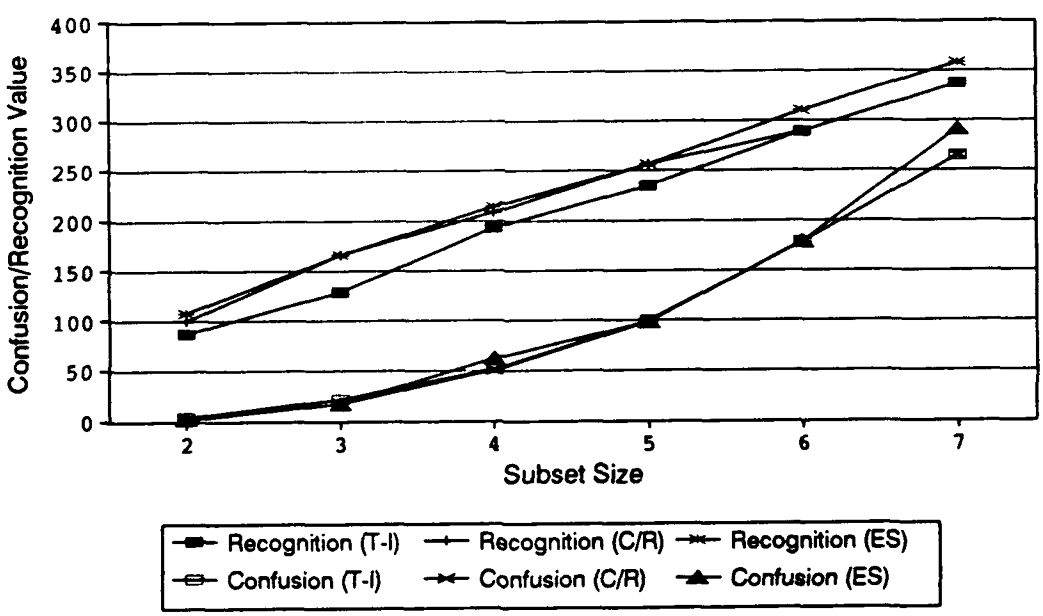

Figure 14 Pollack3 Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	f, r	6	46	0.468
3	f, r, y	37	65	0.427
4	f, r, hw, y	88	92	0.324
5	f, h, r, hw, y	148	137	0.399
6	f, h, r, hw, y, #	261	169	0.341
7	f, h, r, w, hw, y, #	401	197	0.330
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	f, y	6	47	0.472
3	f, r, hw	36	73	0.377
4	f, h, r, hw	84	118	0.418
5	f, h, r, hw, y	148	137	0.399
6	f, h, r, hw, y, #	261	169	0.341
7	f, h, l, r, hw, y, #	396	203	0.298
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	f, y	6	47	0.472
3	h, r, hw	37	90	0.455
4	f, h, r, hw	84	118	0.418
5	f, h, r, hw, y	148	137	0.399
6	f, h, r, hw, y, #	261	169	0.341
7	f, h, r, w, hw, y, #	401	197	0.330

Figure 15 Comparison of Model Results for Pollack4 Data Set

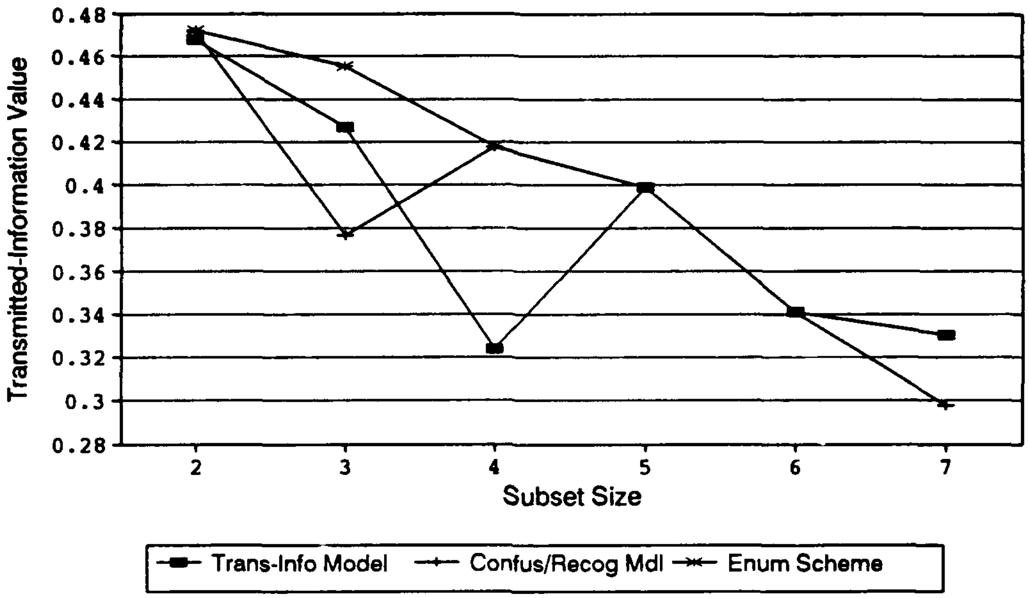

Figure 16 Pollack4 Data Set: Transmitted-Information

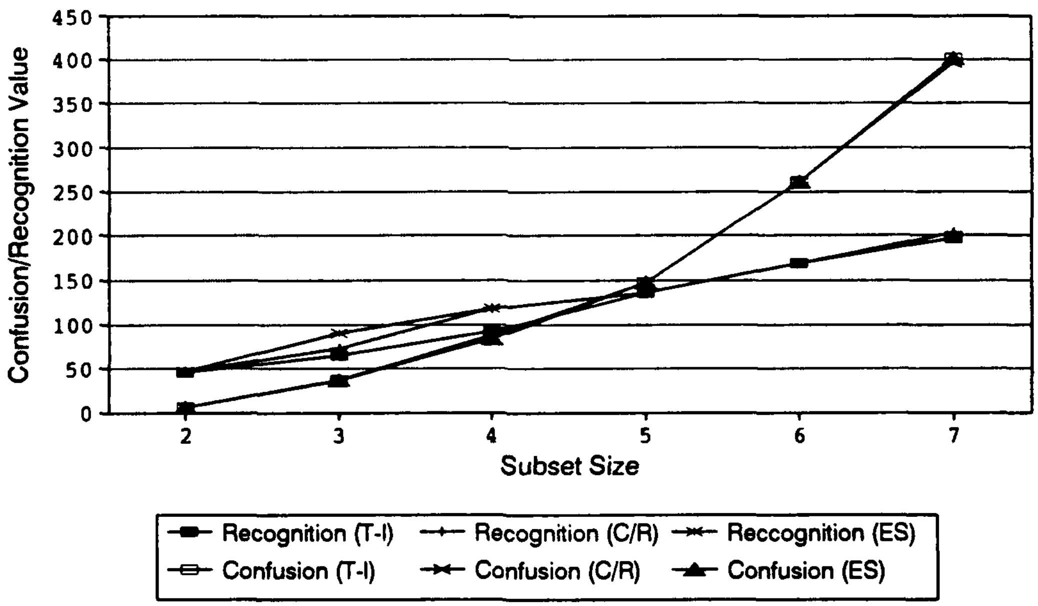

Figure 17 Pollack4 Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	1, 4	0.0	188.9	1.000
3	1, 3, 4	0.0	282.8	1.585
4	1, 3, 4, 5	0.0	375.3	2.000
5	1, 3, 4, 5, 7	0.0	467.3	2.322
6	1, 2, 3, 4, 5, 7	0.0	558.0	2.585
7	1, 2, 3, 4, 5, 7, 8	0.0	648.5	2.807
8	0, 1, 2, 3, 4, 5, 7, 8	5.6	735.0	2.958
9	0, 1, 2, 3, 4, 5, 6, 7, 8	16.0	820.7	3.064
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	1, 4	0.0	188.9	1.000
3	1, 3, 4	0.0	282.8	1.585
4	1, 3, 4, 5	0.0	375.3	2.000
5	1, 3, 4, 5, 7	0.0	467.3	2.322
6	1, 2, 3, 4, 5, 7	0.0	558.0	2.585
7	1, 2, 3, 4, 5, 7, 8	0.0	648.5	2.807
8	0, 1, 2, 3, 4, 5, 7, 8	5.6	735.0	2.958
9	0, 1, 2, 3, 4, 5, 6, 7, 8	16.0	820.7	3.064
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	1, 4	0.0	188.9	1.000
3	1, 3, 4	0.0	282.8	1.585
4	1, 3, 4, 5	0.0	375.3	2.000
5	1, 3, 4, 5, 7	0.0	467.3	2.322
6	1, 2, 3, 4, 5, 7	0.0	558.0	2.585
7	1, 2, 3, 4, 5, 7, 8	0.0	648.5	2.807
8	0, 1, 2, 3, 4, 5, 7, 8	5.6	735.0	2.958
9	0, 1, 2, 3, 4, 5, 6, 7, 8	16.0	820.7	3.064

Figure 18 Comparison of Model Results for Wilpon10 Data Set

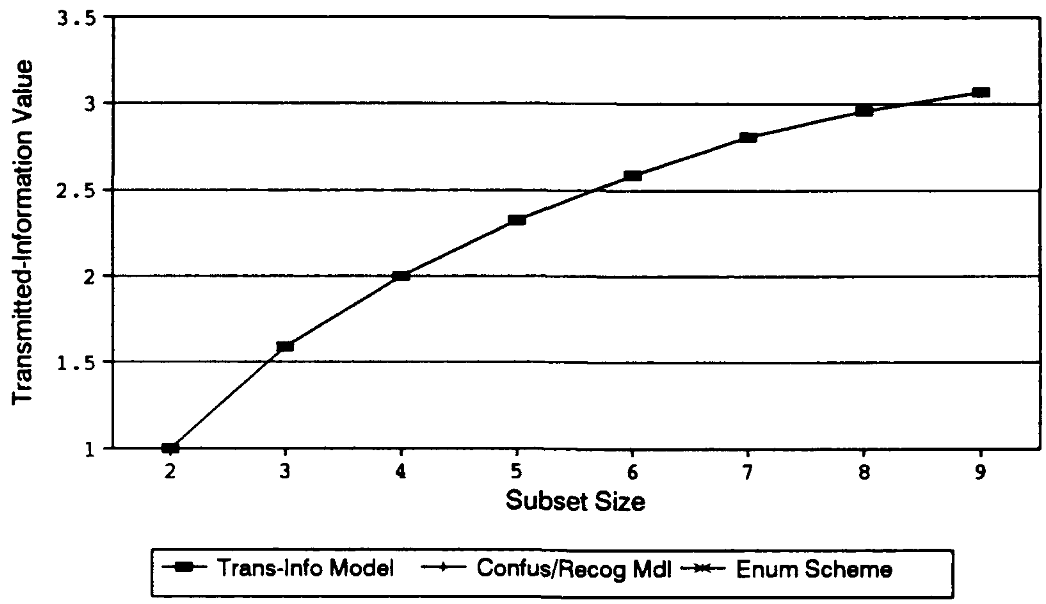

Figure 19 Wilpon10 Data Set: Transmitted-Information

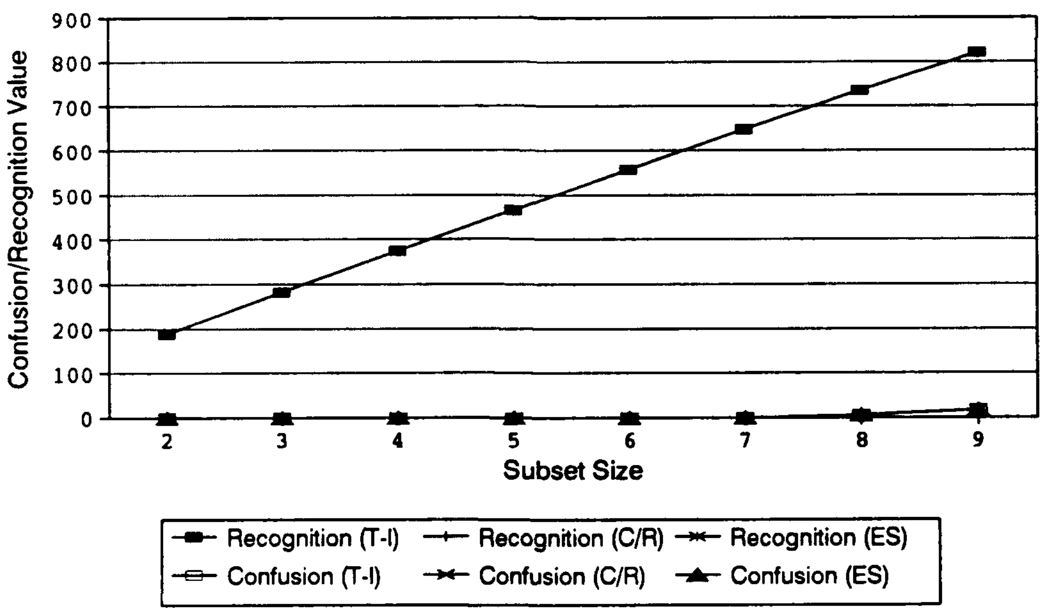

Figure 20 Wilpon10 Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	3, 4	0.0	186.7	1.000
3	3, 4, 7	0.0	275.1	1.584
4	3, 4, 5, 7	0.0	362.9	1.999
5	1, 3, 4, 5, 7	5.3	451.1	2.256
6	1, 2, 3, 4, 5, 7	10.7	529.3	2.473
7	1, 2, 3, 4, 5, 7, 8	18.2	610.5	2.650
8	0, 1, 2, 3, 4, 5, 7, 8	43.5	680.1	2.698
9	0, 1, 2, 3, 4, 5, 7, 8, 9	70.9	755.0	2.740
Subset size	Confusion/Recognition Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	3, 4	0.0	186.7	1.000
3	3, 4, 7	0.0	275.1	1.584
4	3, 4, 5, 7	0.0	362.9	1.999
5	0, 1, 2, 3, 5	4.6	415.1	2.256
6	1, 2, 3, 4, 5, 7	10.7	529.3	2.473
7	1, 2, 3, 4, 5, 7, 8	18.2	610.5	2.650
8	0, 1, 2, 3, 4, 5, 7, 8	43.5	680.1	2.698
9	0, 1, 2, 3, 4, 5, 7, 8, 9	70.9	755.0	2.740
Subset size	Enumeration Scheme Selected Subsets	Confusion/ Recognition		Transmitted information
2	1, 7	0.0	176.6	1.000
3	1, 5, 7	0.0	264.4	1.585
4	1, 3, 5, 7	0.0	355.7	2.000
5	0, 1, 2, 3, 5	4.6	415.1	2.256
6	1, 2, 3, 4, 5, 7	10.7	529.3	2.473
7	1, 2, 3, 4, 5, 7, 8	18.2	610.5	2.650
8	0, 1, 2, 3, 4, 5, 7, 8	43.5	680.1	2.698
9	0, 1, 2, 3, 4, 5, 7, 8, 9	70.9	755.0	2.740

Figure 21 Comparison of Model Results for Wilpon7A Data Set

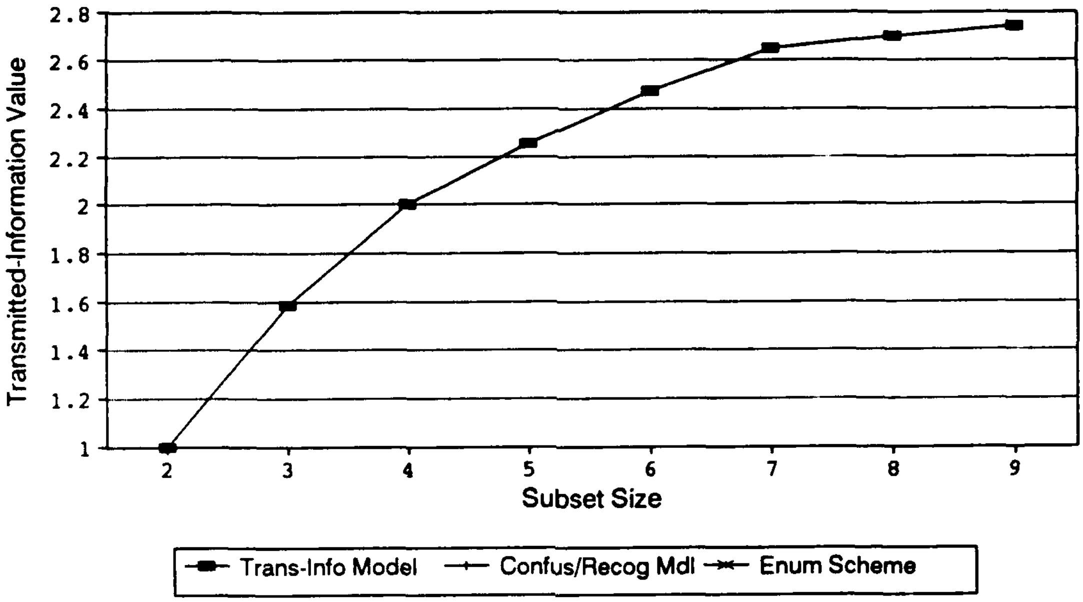

Figure 22 Wilpon7A Data Set: Transmitted-Information

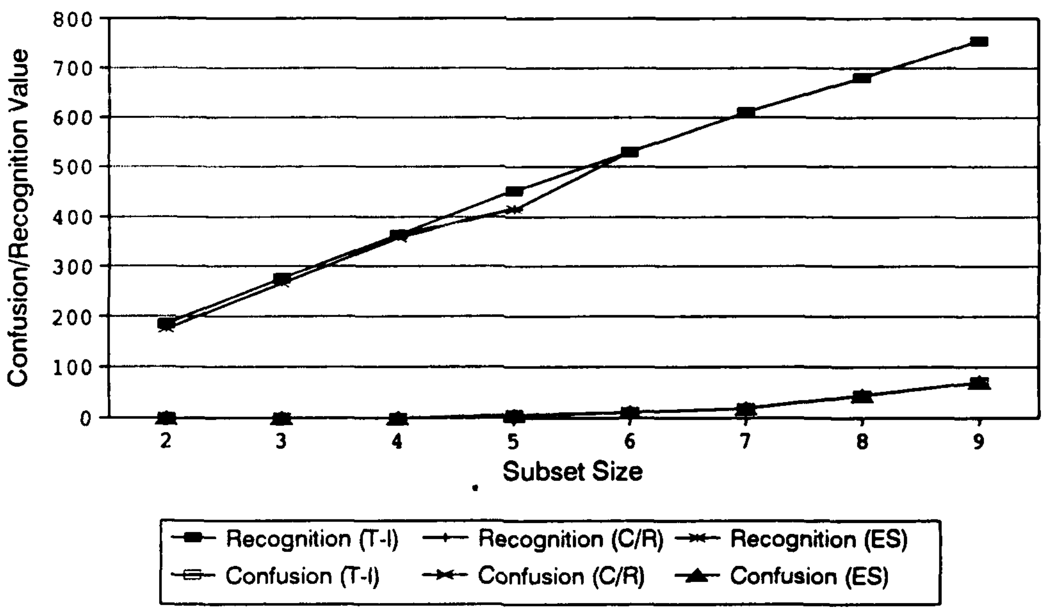

Figure 23 Wilpon7A Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/	Recognition	Transmitted information
2	1, 3	0.0	189.9	1.000
3	1, 3, 8	0.0	280.2	1.585
4	1, 3, 7, 8	0.0	365.7	1.999
5	3, 4, 5, 7, 8	0.0	449.1	2.320
6	0, 1, 3, 5, 7, 8	6.8	516.0	2.506
7	0, 1, 2, 3, 5, 7, 8	19.4	593.8	2.638
8	0, 1, 2, 3, 5, 6, 7, 8	37.1	676.2	2.720
9	0, 1, 2, 3, 4, 5, 6, 7, 8	76.2	769.8	2.752
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/**	**Recognition**	**Transmitted information**
2	1, 3	0.0	189.9	1.000
3	1, 3, 8	0.0	280.2	1.585
4	1, 3, 7, 8	0.0	365.7	1.999
5	3, 4, 5, 7, 8	0.0	449.1	2.320
6	0, 1, 3, 5, 7, 8	6.8	516.0	2.506
7	0, 1, 2, 3, 5, 7, 8	19.4	593.8	2.638
8	0, 1, 2, 3, 5, 6, 7, 8	37.1	676.2	2.720
9	0, 1, 2, 3, 5, 6, 7, 8, 9	64.5	755.2	2.745
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/**	**Recognition**	**Transmitted information**
2	2, 9	0.0	156.8	1.000
3	3, 4, 8	0.0	279.6	1.585
4	3, 4, 7, 8	0.0	365.1	1.999
5	3, 4, 5, 7, 8	0.0	449.1	2.320
6	0, 1, 3, 5, 7, 8	6.8	516.0	2.506
7	0, 1, 2, 3, 5, 7, 8	19.4	593.8	2.638
8	0, 1, 2, 3, 5, 6, 7, 8	37.1	676.2	2.720
9	0, 1, 2, 3, 4, 5, 6, 7, 8	76.2	769.8	2.752

Figure 24 Comparison of Model Results for Wilpon7B Data Set

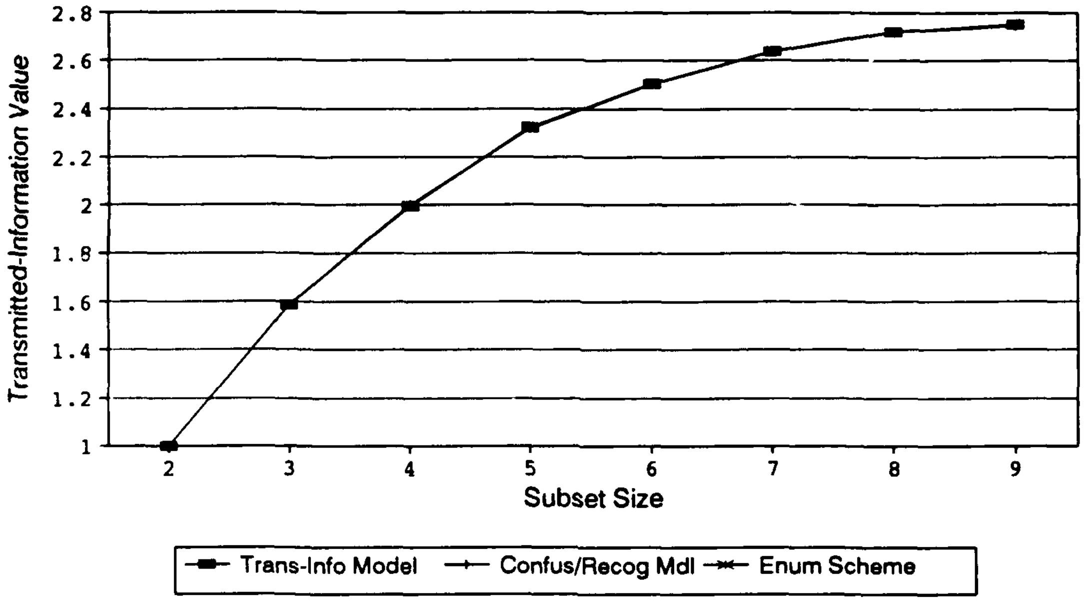

Figure 25 Wilpon7B Data Set: Transmitted-Information

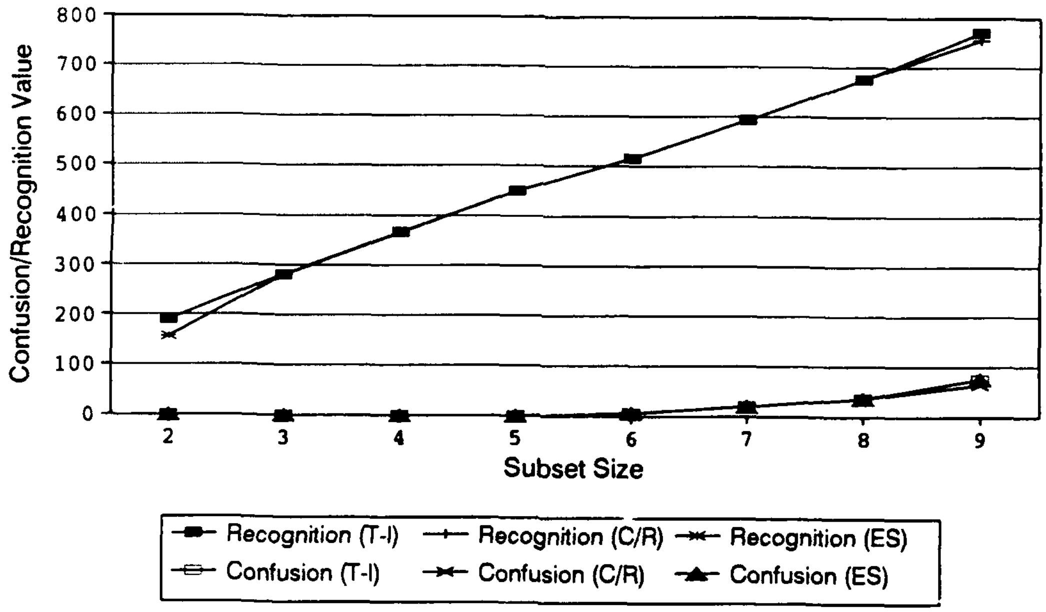

Figure 26 Wilpon7B Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	0, 4	0	200	1.000
3	0, 2, 4	0	299	1.585
4	0, 2, 3, 4	0	398	2.000
5	0, 1, 2, 3, 4	0	496	2.322
6	0, 1, 3, 4, 6, 7	0	578	2.583
7	0, 1, 2, 3, 4, 6, 7	13	677	2.709
8	0, 1, 2, 3, 4, 5, 6, 7	26	752	2.812
9	0, 1, 2, 3, 4, 5, 6, 7, 8	39	839	2.909
Subset size	Confusion/Recognition Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	0, 4	0	200	1.000
3	0, 3, 4	0	299	1.585
4	0, 2, 3, 4	0	398	2.000
5	0, 1, 2, 3, 4	0	496	2.322
6	0, 1, 3, 4, 6, 7	0	578	2.583
7	0, 1, 3, 4, 6, 7, 8	10	665	2.712
8	0, 1, 3, 4, 5, 6, 7, 8	23	740	2.814
9	0, 1, 2, 3, 4, 5, 6, 7, 8	39	839	2.909
Subset size	Enumeration Scheme Selected Subsets	Confusion/ Recognition		Transmitted information
2	0, 4	0	200	1.000
3	0, 2, 4	0	299	1.585
4	0, 2, 3, 4	0	398	2.000
5	0, 1, 2, 3, 4	0	496	2.322
6	0, 1, 3, 4, 6, 7	0	578	2.583
7	0, 1, 3, 4, 6, 7, 8	10	665	2.712
8	0, 1, 3, 4, 5, 6, 7, 8	23	740	2.814
9	0, 1, 2, 3, 4, 5, 6, 7, 8	39	839	2.909

Figure 27 Comparison of Model Results for Wilpon7C Data Set

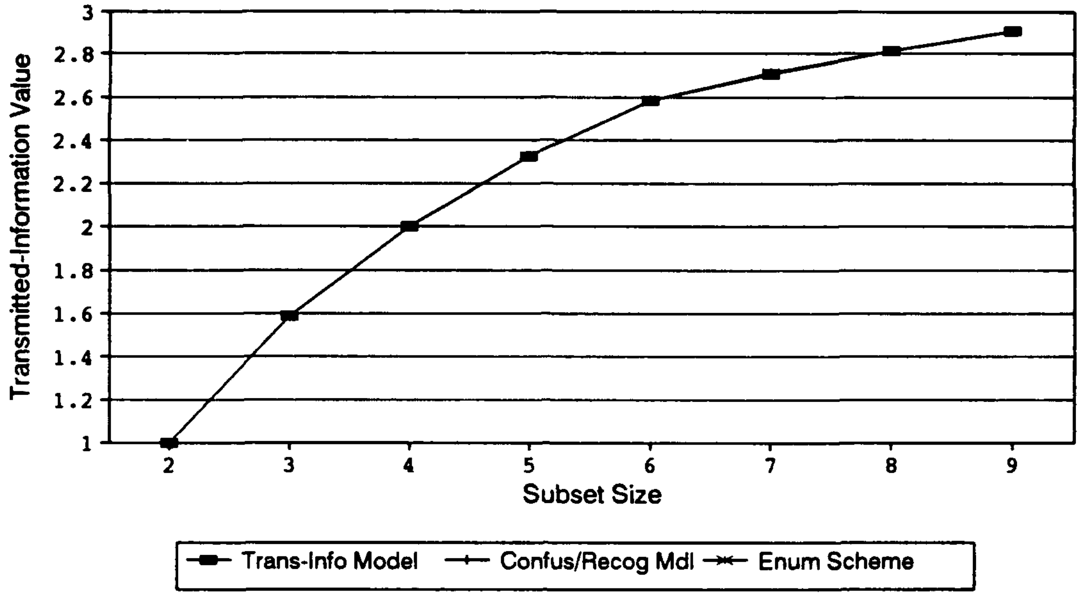

Figure 28 Wilpon7C Data Set: Transmitted-Information

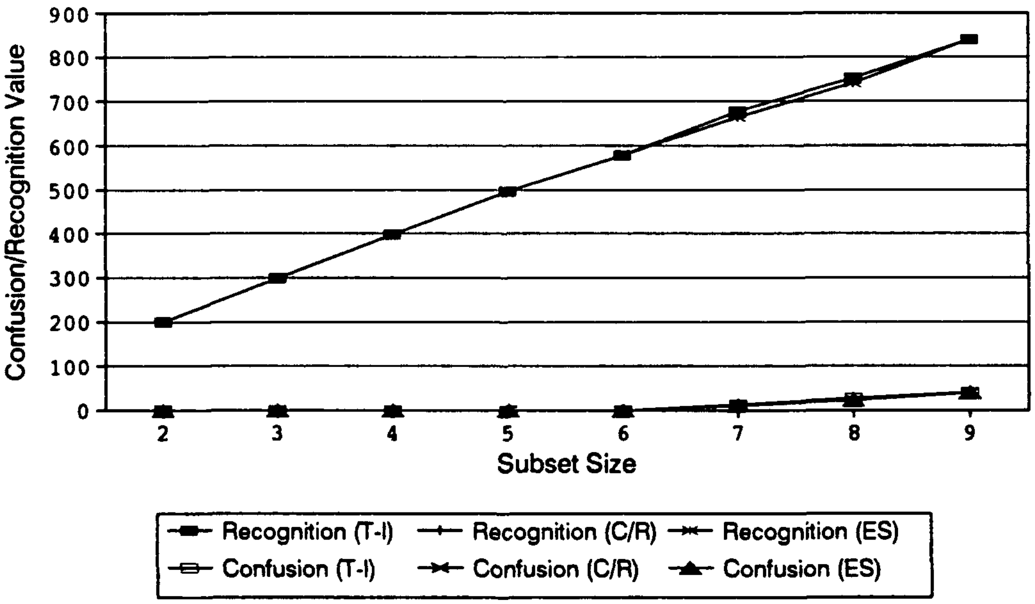

Figure 29 Wilpon7C Data Set: Confusion/Recognition

Subset size	*Transmitted-Information Model Selected Subsets*	*Confusion/*	*Recognition*	*Transmitted information*
2	2, 9	0.0	91.1	0.998
3	2, 4, 9	0.0	141.7	1.582
4	2, 3, 4, 9	3.3	215.9	1.879
5	1, 2, 3, 4, 9	21.1	262.2	1.952
6	1, 2, 3, 4, 8, 9	55.9	341.4	1.929
7	0, 1, 2, 3, 4, 8, 9	110.6	399.8	1.914
8	0, 1, 2, 3, 4, 5, 8, 9	173.1	479.4	1.928
9	0, 1, 2, 3, 4, 5, 6, 8, 9	242.8	542.0	1.889
Subset size	*Confusion/Recognition Model Selected Subsets*	*Confusion/*	*Recognition*	*Transmitted information*
2	5, 8	0.0	158.8	1.000
3	0, 3, 9	0.0	175.8	1.551
4	2, 3, 4, 9	3.3	215.9	1.879
5	1, 2, 3, 4, 9	21.1	262.2	1.952
6	1, 2, 3, 4, 6, 9	53.8	324.8	1.922
7	0, 1, 2, 3, 4, 6, 9	106.7	383.2	1.909
8	0, 1, 2, 3, 4, 6, 8, 9	172.9	462.4	1.867
9	0, 1, 2, 3, 4, 5, 6, 8, 9	242.8	542.0	1.889
Subset size	*Enumeration Scheme Selected Subsets*	*Confusion/*	*Recognition*	*Transmitted information*
2	5, 8	0.0	158.8	1.000
3	2, 4, 9	0.0	141.7	1.582
4	2, 3, 4, 9	3.3	215.9	1.879
5	1, 2, 3, 4, 9	21.1	262.2	1.952
6	1, 2, 3, 4, 8, 9	55.9	341.4	1.929
7	0, 2, 3, 4, 5, 8, 9	110.6	433.1	1.919
8	0, 1, 2, 3, 4, 5, 8, 9	173.1	479.4	1.928
9	0, 1, 2, 3, 4, 5, 6, 8, 9	242.8	542.0	1.889

Figure 30 Comparison of Model Results for Wilpon8A Data Set

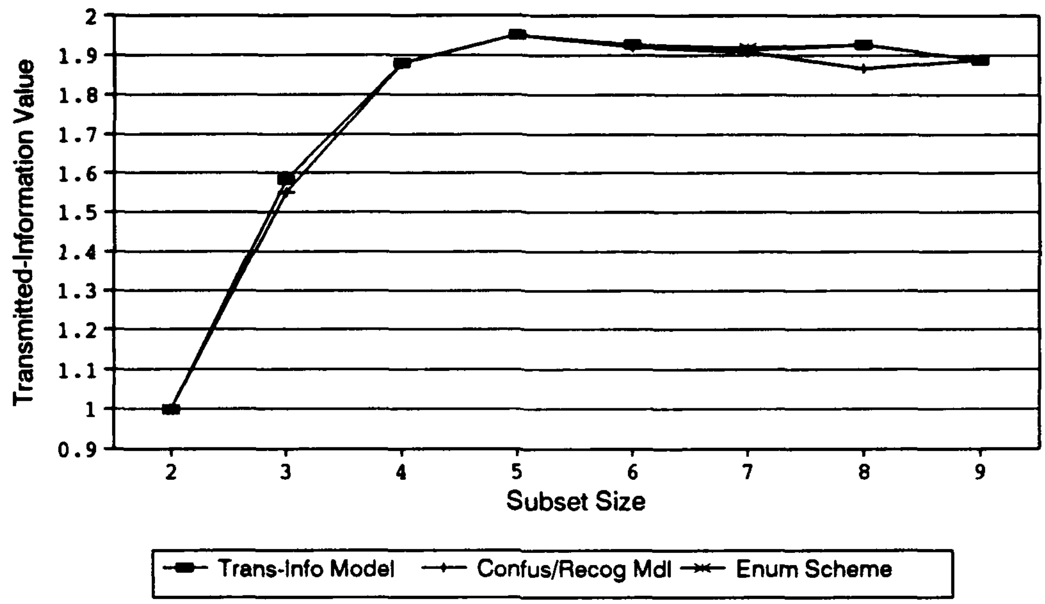

Figure 31 Wilpon8A Data Set: Transmitted-Information

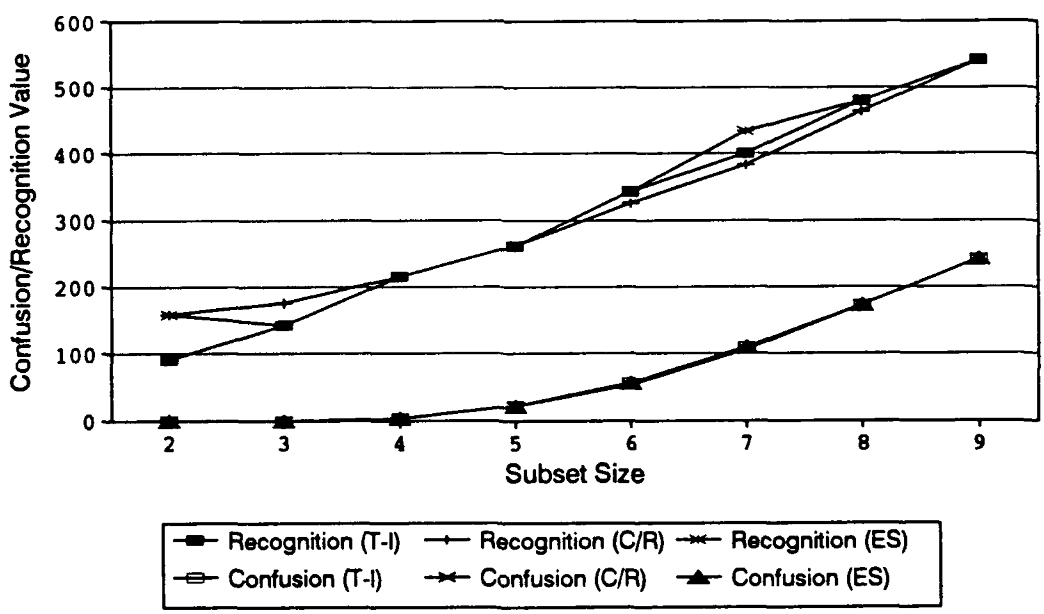

Figure 32 Wilpon8A Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	4, 8	0.0	144.0	0.923
3	2, 4, 8	0.0	230.4	1.531
4	2, 4, 8, 9	0.0	294.7	1.954
5	1, 2, 3, 4, 6	0.0	381.5	2.292
6	1, 2, 3, 4, 6, 7	15.2	456.3	2.400
7	1, 2, 3, 4, 6, 7, 8	35.4	551.6	2.455
8	1, 2, 3, 4, 5, 6, 7, 8	60.8	632.0	2.520
9	0, 1, 2, 3, 4, 5, 6, 7, 8	113.2	716.3	2.527
Subset size	Confusion/Recognition Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	2, 8	0.0	181.7	0.998
3	2, 5, 8	0.0	262.1	1.581
4	1, 2, 3, 6	0.0	332.8	1.996
5	1, 2, 3, 4, 6	0.0	381.5	2.292
6	1, 2, 3, 4, 6, 7	15.2	456.3	2.400
7	1, 2, 3, 4, 6, 7, 8	35.4	551.6	2.455
8	1, 2, 3, 4, 5, 6, 7, 8	60.8	632.0	2.520
9	1, 2, 3, 4, 5, 6, 7, 8, 9	110.0	696.3	2.454
Subset size	Enumeration Scheme Selected Subsets	Confusion/ Recognition		Transmitted information
2	2, 6	0.0	172.2	1.000
3	2, 3, 6	0.0	259.9	1.585
4	1, 2, 3, 6	0.0	332.8	1.996
5	1, 2, 3, 4, 6	0.0	381.5	2.292
6	1, 2, 3, 4, 6, 7	15.2	456.3	2.400
7	1, 2, 3, 4, 6, 7, 8	35.4	551.6	2.455
8	1, 2, 3, 4, 5, 6, 7, 8	60.8	632.0	2.520
9	0, 1, 2, 3, 4, 5, 6, 7, 8	113.2	716.3	2.527

Figure 33 Comparison of Model Results for Wilpon8B Data Set

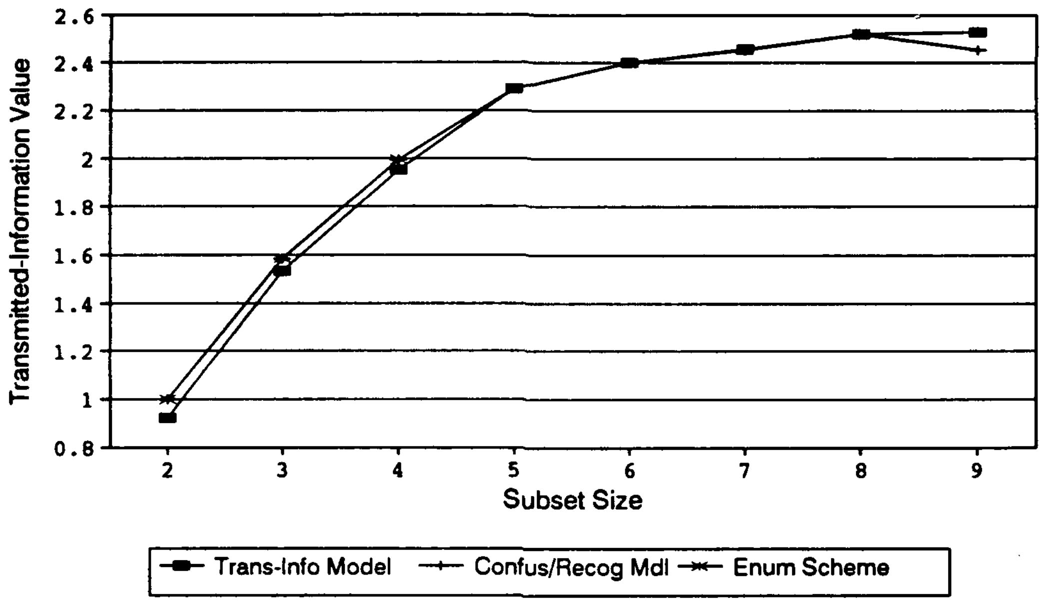

Figure 34 Wilpon8B Data Set: Transmitted-Information

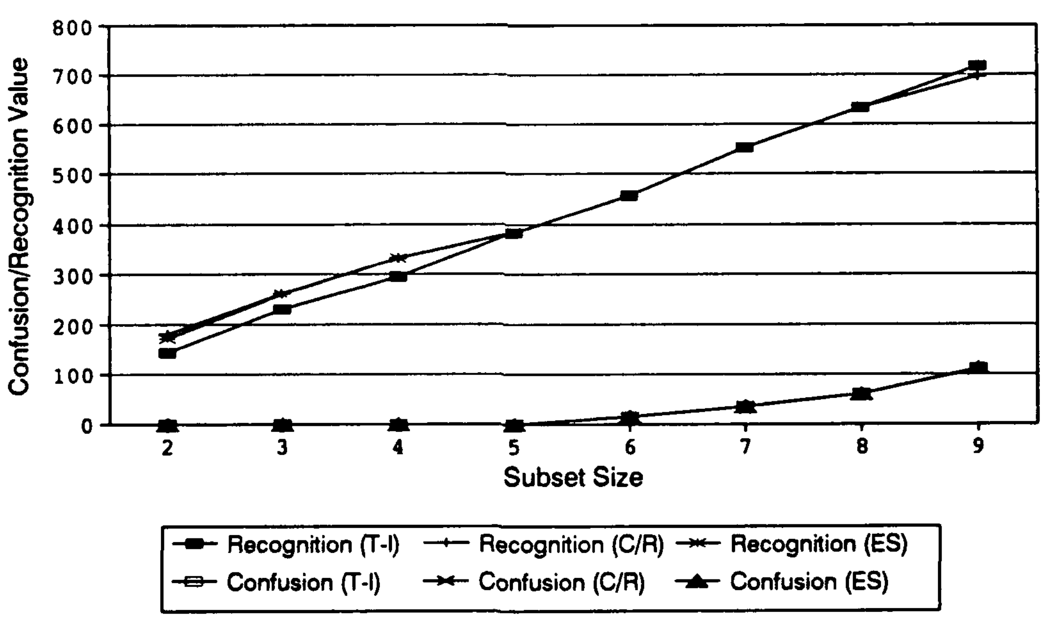

Figure 35 Wilpon8B Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition	Transmitted Information
2	0, 2	0 200	1.000
3	0, 1, 2	0 300	1.585
4	0, 2, 3, 7	0 400	2.000
5	0, 1, 2, 3, 7	0 500	2.322
6	0, 1, 2, 3, 4, 7	0 599	2.585
7	0, 1, 2, 3, 4, 7, 9	0 698	2.807
8	0, 1, 2, 3, 4, 7, 8, 9	0 796	3.000
9	0, 1, 2, 3, 4, 5, 7, 8, 9	0 893	3.170

Subset size	Confusion/Recognition Model Selected Subsets	Confusion/ Recognition	Transmitted information
2	0, 7	0 200	1.000
3	0, 3, 7	0 300	1.585
4	0, 1, 3, 7	0 400	2.000
5	0, 1, 2, 3, 7	0 500	2.322
6	0, 1, 2, 3, 4, 7	0 599	2.585
7	0, 1, 2, 3, 4, 7, 9	0 698	2.807
8	0, 1, 2, 3, 4, 7, 8, 9	0 796	3.000
9	0, 1, 2, 3, 4, 5, 7, 8, 9	0 893	3.170

Subset size	Enumeration Scheme Selected Subsets	Confusion/ Recognition	Transmitted information
2	0, 1	0 200	1.000
3	0, 1, 2	0 300	1.585
4	0, 1, 2, 3	0 400	2.000
5	0, 1, 2, 3, 7	0 500	2.322
6	0, 1, 2, 3, 4, 7	0 599	2.585
7	0, 1, 2, 3, 4, 7, 9	0 698	2.807
8	0, 1, 2, 3, 4, 7, 8, 9	0 796	3.000
9	0, 1, 2, 3, 4, 5, 7, 8, 9	0 893	3.170

Figure 36 Comparison of Model Results for Wilpon8C Data Set

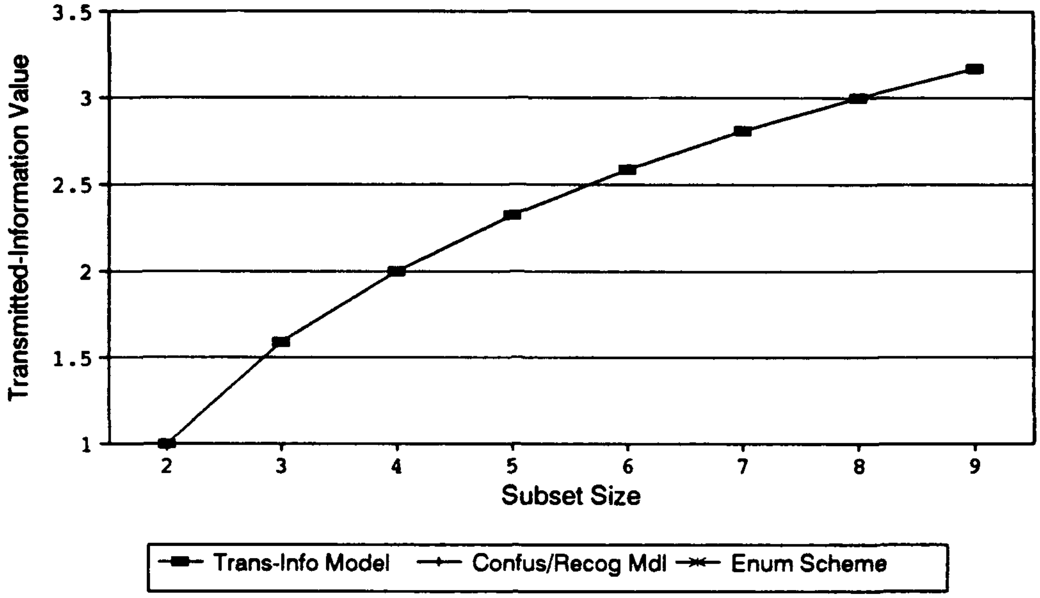

Figure 37 Wilpon8C Data Set: Transmitted-Information

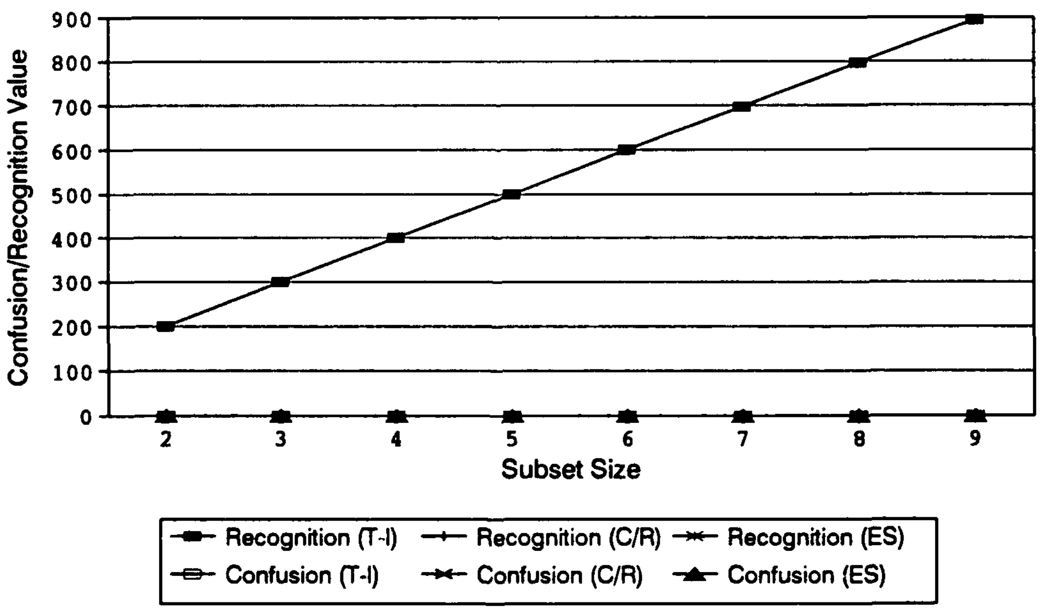

Figure 38 Wilpon8C Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/	Recognition	Transmitted information
2	4, 5	0.0	176.2	1.000
3	3, 4, 5	0.0	260.8	1.585
4	0, 1, 3, 9	0.0	284.7	1.987
5	0, 1, 3, 8, 9	10.0	357.2	2.183
6	0, 1, 3, 4, 8, 9	34.1	445.7	2.233
7	0, 1, 3, 4, 5, 8, 9	62.2	533.4	2.323
8	0, 1, 3, 4, 5, 7, 8, 9	91.8	618.3	2.384
9	0, 1, 2, 3, 4, 5, 7, 8, 9	125.3	685.1	2.429
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/**	**Recognition**	**Transmitted information**
2	4, 5	0.0	176.2	1.000
3	3, 4, 5	0.0	260.8	1.585
4	0, 1, 3, 9	0.0	284.7	1.986
5	0, 1, 3, 8, 9	10.0	357.2	2.183
6	0, 1, 3, 5, 7, 8	32.7	469.7	2.242
7	0, 1, 3, 4, 5, 7, 8	56.8	558.2	2.340
8	0, 1, 2, 3, 4, 5, 7, 8	90.3	625.0	2.370
9	0, 1, 2, 3, 4, 5, 7, 8, 9	125.3	685.1	2.429
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/**	**Recognition**	**Transmitted information**
2	3, 7	0.0	169.5	1.000
3	3, 4, 5	0.0	260.8	1.585
4	0, 1, 8, 9	0.0	272.6	1.994
5	0, 1, 3, 8, 9	10.0	357.2	2.183
6	0, 1, 3, 5, 7, 8	32.7	469.7	2.242
7	0, 1, 3, 4, 5, 7, 8	56.8	558.2	2.340
8	0, 1, 2, 3, 4, 5, 7, 8	91.5	600.2	2.385
9	0, 1, 2, 3, 4, 5, 7, 8, 9	125.3	685.1	2.429

Figure 39 Comparison of Model Results for Wilpon9A Data Set

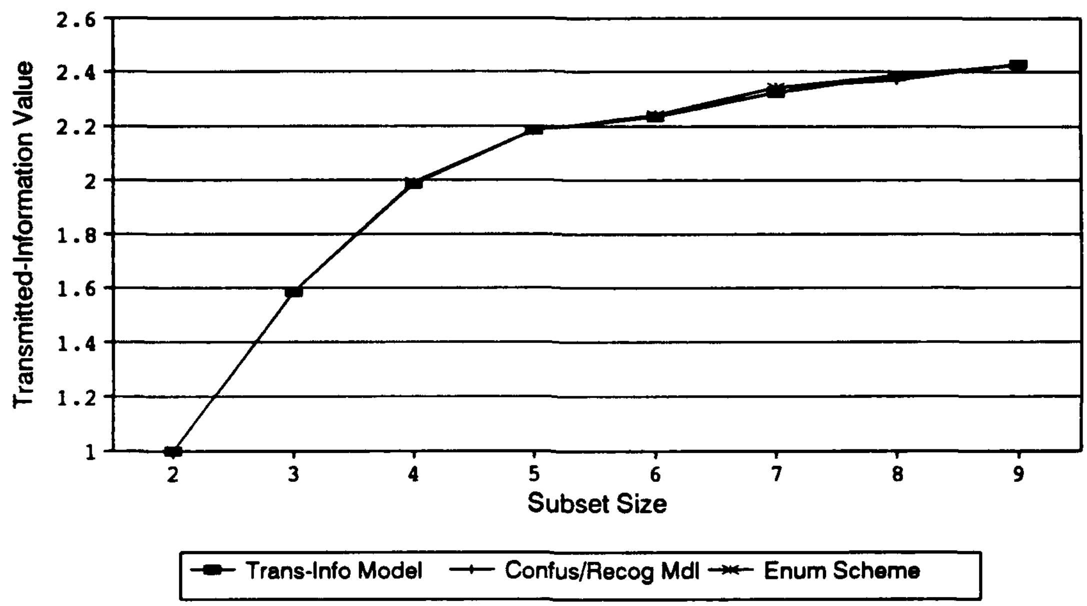

Figure 40 Wilpon9A Data Set: Transmitted-Information

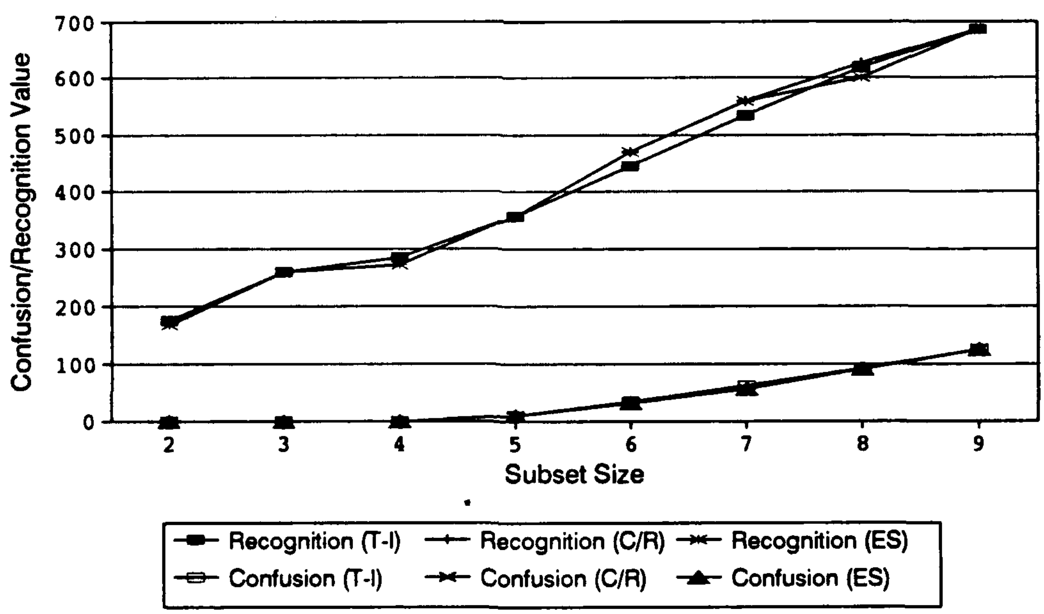

Figure 41 Wilpon9A Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition	Transmitted information
2	1, 8	0.0 190.5	1.000
3	1, 2, 8	0.0 285.5	1.585
4	1, 2, 3, 8	0.0 379.7	2.000
5	1, 2, 3, 4, 8	0.0 473.1	2.322
6	1, 2, 3, 4, 7, 8	0.0 566.4	2.585
7	1, 2, 3, 4, 5, 7, 8	0.0 659.5	2.807
8	0, 1, 2, 3, 4, 5, 7, 8	3.8 750.4	2.969
9	0, 1, 2, 3, 4, 5, 6, 7, 8	11.2 837.4	3.089
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/ Recognition**	**Transmitted information**
2	1, 8	0.0 190.5	1.000
3	1, 2, 8	0.0 285.5	1.585
4	1, 2, 3, 8	0.0 379.7	2.000
5	1, 2, 3, 4, 8	0.0 473.1	2.322
6	1, 2, 3, 4, 7, 8	0.0 566.4	2.585
7	1, 2, 3, 4, 5, 7, 8	0.0 659.5	2.807
8	0, 1, 2, 3, 4, 5, 7, 8	3.8 750.4	2.969
9	0, 1, 2, 3, 4, 5, 6, 7, 8	11.2 837.4	3.089
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/ Recognition**	**Transmitted information**
2	1, 8	0.0 190.5	1.000
3	1, 2, 8	0.0 285.5	1.585
4	1, 2, 3, 8	0.0 379.7	2.000
5	1, 2, 3, 4, 8	0.0 473.1	2.322
6	1, 2, 3, 4, 7, 8	0.0 566.4	2.585
7	1, 2, 3, 4, 5, 7, 8	0.0 659.5	2.807
8	0, 1, 2, 3, 4, 5, 7, 8	3.8 750.4	2.969
9	0, 1, 2, 3, 4, 5, 6, 7, 8	11.2 837.4	3.089

Figure 42 Comparison of Model Results for Wilpon9B Data Set

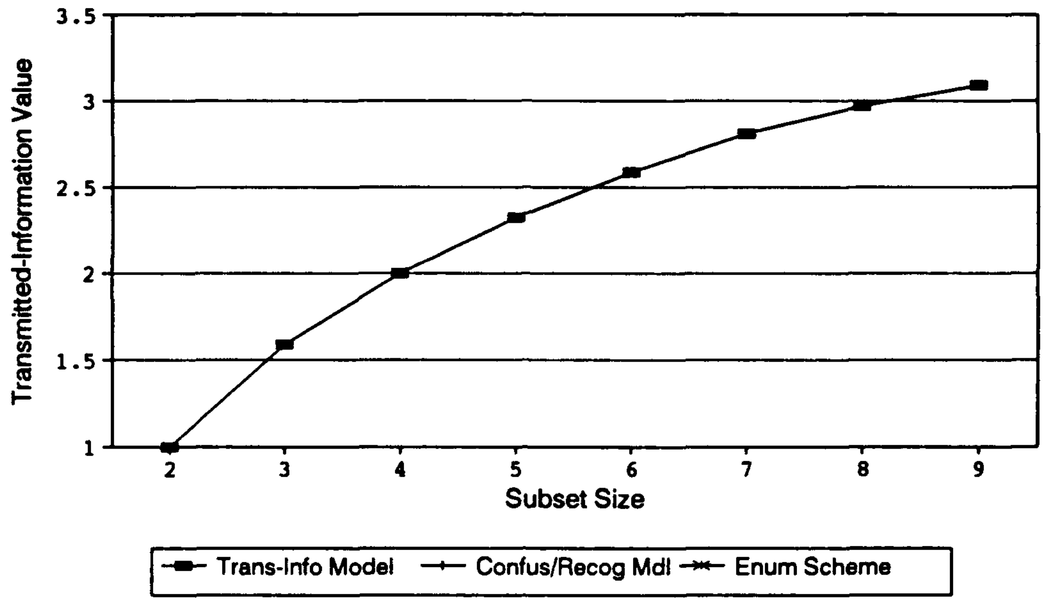

Figure 43 Wilpon9B Data Set: Transmitted-Information

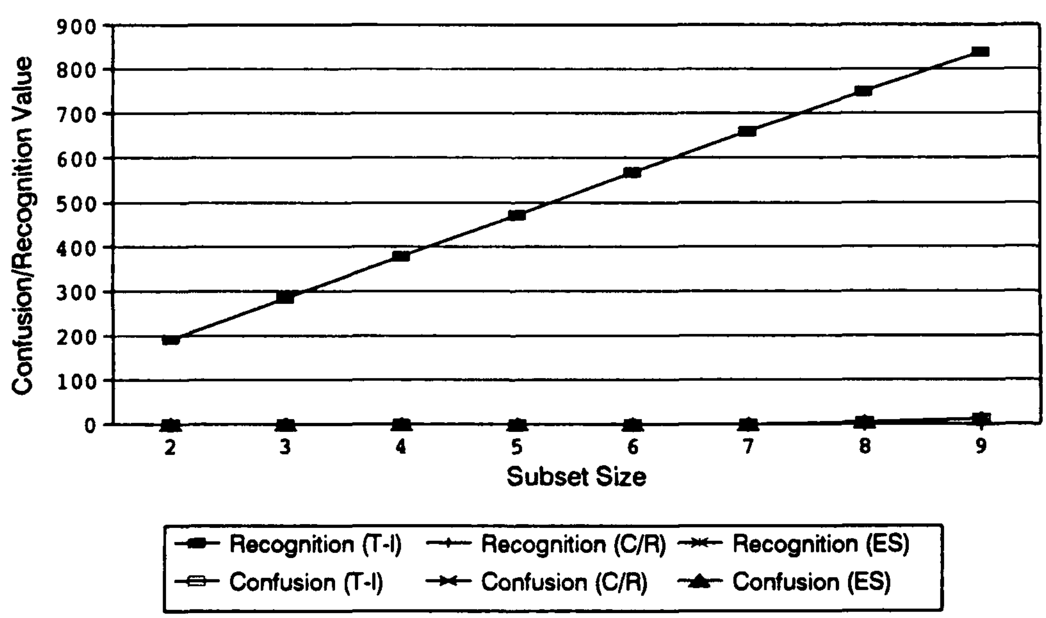

Figure 44 Wilpon9B Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	1, 2	0	196	1.000
3	1, 2, 3	0	294	1.585
4	1, 3, 4, 7	0	384	1.999
5	0, 1, 3, 5, 6	0	446	2.313
6	0, 1, 3, 5, 6, 8	3	540	2.541
7	0, 1, 3, 4, 6, 7, 8	12	656	2.699
8	0, 1, 3, 4, 6, 7, 8, 9	19	746	2.857
9	0, 1, 2, 3, 4, 6, 7, 8, 9	42	844	2.902
Subset size	**Confusion/Recognition Model Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	1, 3	0	196	1.000
3	1, 2, 3	0	294	1.585
4	1, 3, 4, 6	0	384	1.999
5	0, 1, 3, 5, 6	0	446	2.313
6	0, 1, 3, 5, 6, 8	3	540	2.541
7	0, 1, 3, 5, 6, 7, 8	11	631	2.697
8	0, 1, 3, 4, 6, 7, 8, 9	19	746	2.857
9	0, 1, 2, 3, 4, 6, 7, 8, 9	42	844	2.902
Subset size	**Enumeration Scheme Selected Subsets**	**Confusion/ Recognition**		**Transmitted information**
2	1, 2	0	196	1.000
3	1, 2, 3	0	294	1.585
4	0, 6, 8, 9	0	362	1.999
5	0, 1, 5, 6, 8	0	442	2.314
6	0, 1, 3, 5, 6, 8	3	540	2.541
7	0, 1, 3, 4, 6, 7, 8	12	656	2.699
8	0, 1, 3, 4, 6, 7, 8, 9	19	746	2.857
9	0, 1, 2, 3, 4, 6, 7, 8, 9	44	818	2.904

Figure 45 Comparison of Model Results for Wilpon9C Data Set

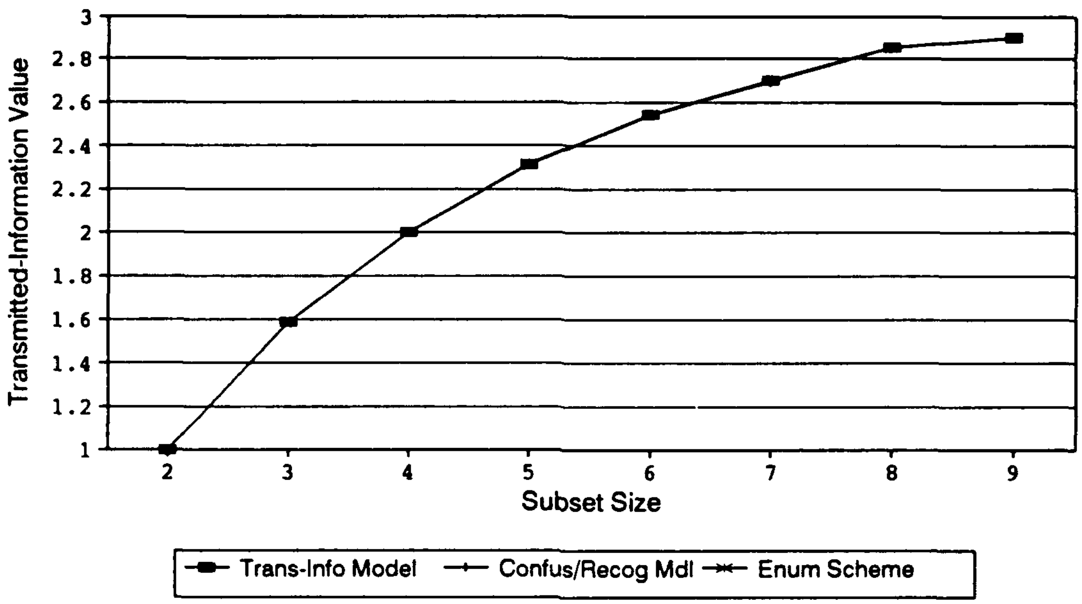

Figure 46 Wilpon9C Data Set: Transmitted-Information

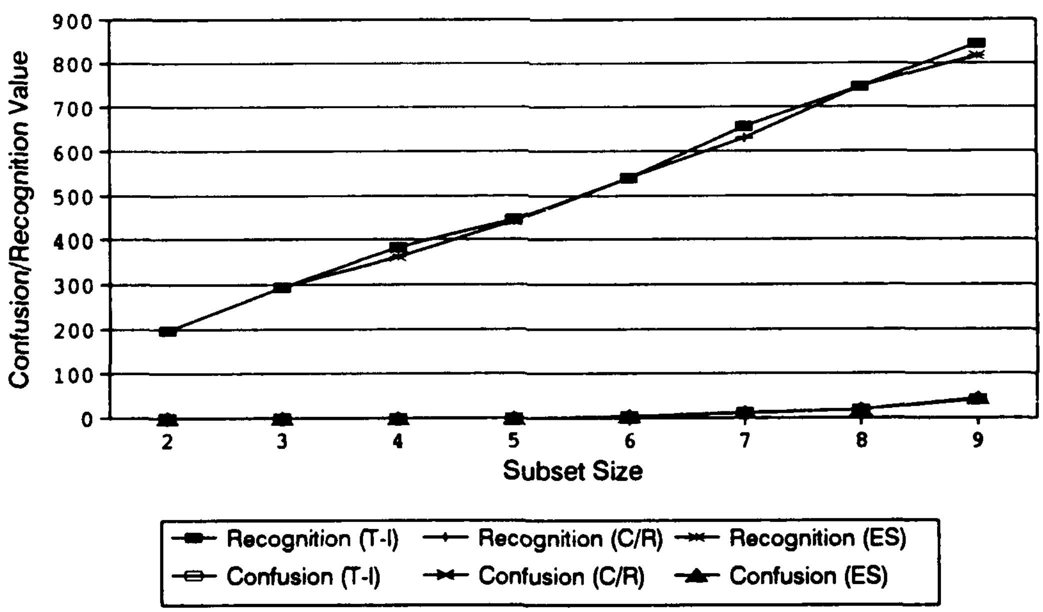

Figure 47 Wilpon9C Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	5, 17	0.000	1.297	0.937
3	5, 7, 14	0.000	2.220	1.546
4	4, 6, 7, 8	0.012	3.458	1.970
5	1, 2, 3, 4, 5	0.048	4.391	2.233
6	1, 2, 4, 5, 6, 7	0.090	5.278	2.446
7	1, 2, 3, 4, 5, 6, 7	0.138	6.147	2.630
8	1, 2, 3, 4, 5, 6, 13, 16	0.210	6.604	2.742
9	1, 2, 4, 5, 6, 10, 11, 13, 16	0.336	7.383	2.821
10	1, 2, 3, 4, 6, 10, 12, 13, 16, 20	0.480	8.074	2.885
11	1, 2, 3, 4, 6, 10, 11, 12, 13, 16, 20	0.672	8.859	2.932
12	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 16	0.971	9.960	2.979
13	1-11, 13, 16	1.229	10.799	2.994
14	1-11, 13, 14, 16	1.712	11.305	2.994
15	1-14, 16	2.036	12.061	2.995
16	1-14, 16, 20	2.437	12.751	2.993
17	1-16, 20	2.815	13.513	2.990
18	1-16, 18, 20	3.264	14.233	2.975
19	1-16, 18-20	3.992	14.792	2.933

Subset size	Confusion/Recognition Model Selected Subsets	Confusion/ Recognition		Transmitted information
2	1, 2	0.000	1.814	1.000
3	1, 2, 3	0.000	2.683	1.585
4	4, 6, 7, 8	0.012	3.458	1.970
5	1, 2, 3, 4, 5	0.048	4.391	2.233
6	1, 2, 4, 5, 6, 7	0.090	5.278	2.446
7	1, 2, 3, 4, 5, 6, 7	0.138	6.147	2.630
8	1, 2, 3, 4, 5, 6, 13, 16	0.210	6.604	2.742
9	1, 2, 4, 5, 6, 10, 11, 13, 16	0.336	7.383	2.821
10	1, 2, 3, 4, 6, 10, 12, 13, 16, 20	0.480	8.074	2.885
11	1, 2, 3, 4, 6, 10, 11, 12, 13, 16, 20	0.672	8.859	2.932
12	1, 2, 3, 4, 6, 8, 10, 11, 12, 13, 16, 20	0.900	9.692	2.959
13	1-4, 6, 8-13, 16, 20	1.152	10.531	2.976
14	1-4, 6, 8-13, 15, 16, 20	1.494	11.293	2.962
15	1-4, 6-13, 15, 16, 20	1.871	12.168	2.981
16	1-13, 15, 16, 20	2.278	12.888	2.990
17	1-13, 15, 16, 18, 20	2.715	13.727	2.975
18	1-16, 18, 20	3.264	14.233	2.975
19	1-18, 20	3.988	14.691	2.902

Subset size	Enumeration Scheme Selected Subsets	Confusion/ Recognition		Transmitted information
2	6, 7	0.000	1.756	1.000
3	1, 2, 3	0.000	2.683	1.585
4	4, 6, 7, 8	0.000	3.458	1.970
5	1, 2, 3, 4, 5	0.048	4.391	2.234
6	1, 2, 4, 5, 6, 7	0.090	5.278	2.446
7	1, 2, 3, 4, 5, 6, 7	0.138	6.147	2.630
8	1, 2, 3, 4, 5, 6, 7, 13	0.222	6.926	2.755
9	1, 2, 3, 4, 5, 6, 7, 8, 13	0.360	7.759	2.836
10	1, 2, 3, 4, 5, 6, 7, 8, 10, 13	0.522	8.622	2.898
11	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13	0.720	9.407	2.942
12	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 16	0.971	9.960	2.979
13	1-11, 13, 16	1.229	10.799	2.994
14	1-11, 13, 16, 20	1.588	11.489	2.996
15	1-13, 16, 20	1.912	12.245	2.997
16	1-14, 16, 20	2.437	12.751	2.993
17	1-16, 20	2.815	13.513	2.990
18	1-16, 18, 20	3.264	14.233	2.981
19	1-16, 18-20	3.992	14.792	2.938

Figure 48 Comparison of Model Results for Bowen Data Set

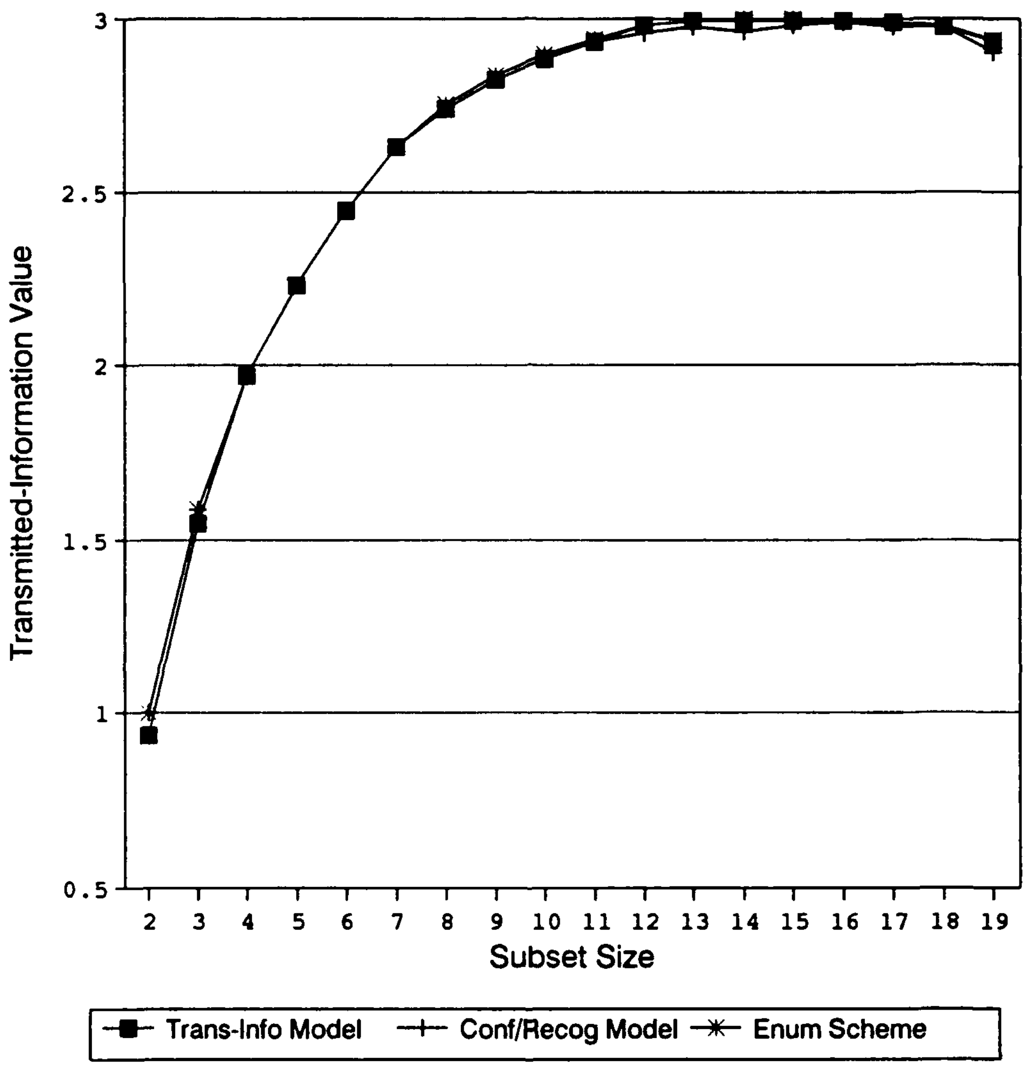

Figure 49 Bowen Data Set: Transmitted-Information

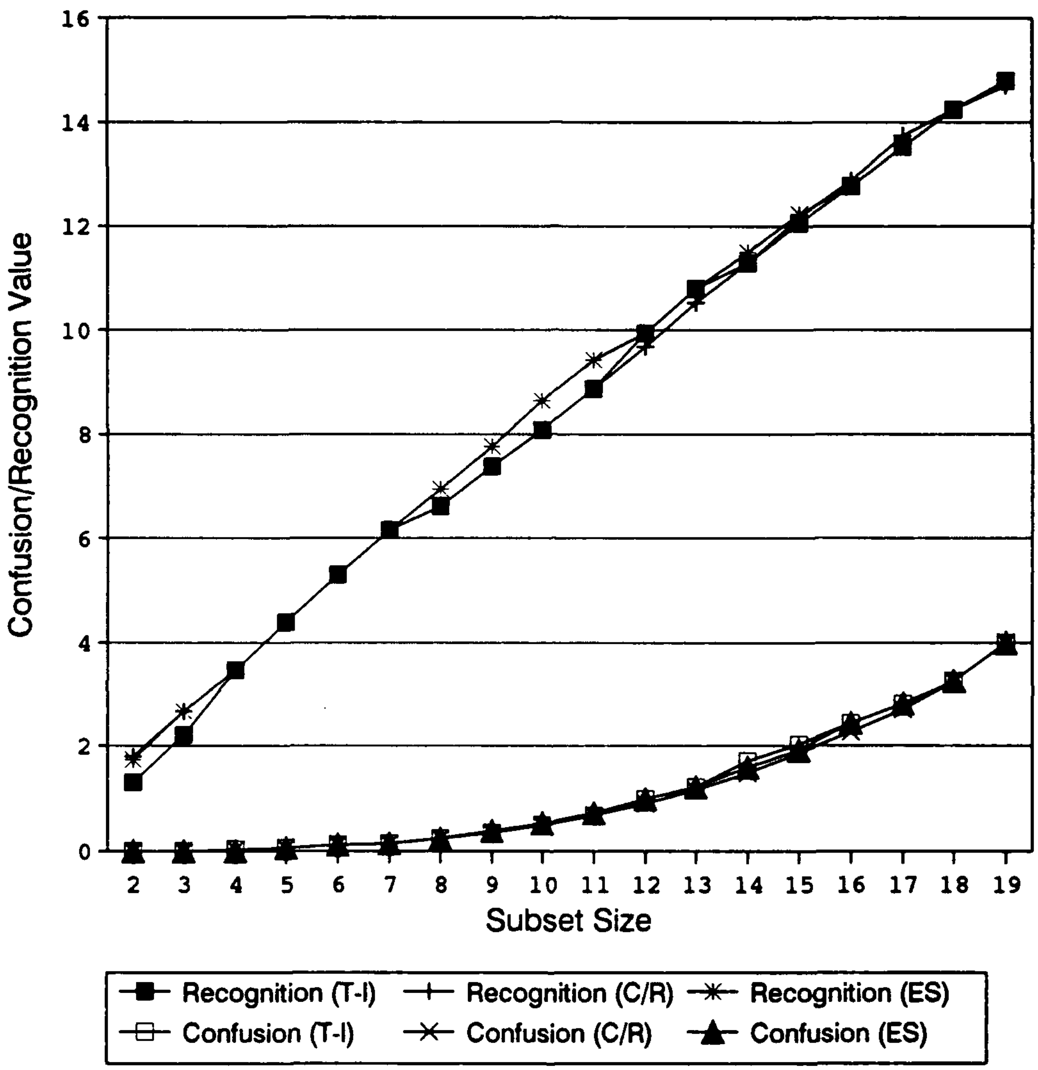

Figure 50 Bowen Data Set: Confusion/Recognition

Subset size	Transmitted-Information Model Selected Subsets	Confusion/ Recognition		Transmitted Information
2	11, 21	5	116	0.903
3	11, 21, 23	5	195	1.511
4	4, 11, 21, 23	5	266	1.945
5	4, 10, 13, 21, 23	0	340	2.283
6	2, 4, 10, 14, 21, 23	0	410	2.553
7	2, 4, 10, 14, 19, 21, 23	0	464	2.774
8	2, 4, 10, 14, 19, 21, 22, 25	1	533	2.957
9	4, 5, 10, 13, 14, 17, 19, 21, 23	2	587	3.114
10	4, 5, 10, 13, 14, 17, 19, 21, 22, 23	5	662	3.239
11	4, 5, 10, 13, 14, 17, 19, 21-24	9	737	3.349
12	1, 4, 8, 11, 13, 17, 18, 20-24	19	805	3.445
13	1, 4, 11, 13, 15, 17-24	25	868	3.519
14	4, 6, 7, 13, 14, 15, 17-24	28	940	3.589
15	4, 6, 7, 13, 14, 15, 17-25	39	1013	3.639
16	2, 4, 6, 7, 13, 14, 15, 17-25	51	1084	3.684
17	1, 2, 4, 6, 7, 13, 14, 15, 17-25	69	1149	3.711
18	1-4, 6, 7, 13, 14, 15, 17-25	97	1196	3.717
19	1-4, 6-8, 13-15, 17-25	129	1256	3.712
20	1-4, 6-9, 13-15, 17-25	170	1319	3.699
21	1-8, 13-25	205	1363	3.670
22	1-9, 13-25	250	1426	3.654
23	1-9, 11, 13-25	300	1463	3.625
24	1-9, 11-25	378	1510	3.600

Figure 51 Results from Trans-Info Model for Moore Data Set

Subset size	Confusion/Recognition Model Selected Subsets	Confusion/	Recognition	Transmitted Information
2	21, 23	0	158	1.000
3	13, 21, 23	0	230	1.584
4	4, 13, 21, 23	0	301	1.998
5	2, 4, 14, 21, 23	0	371	2.320
6	2, 4, 14, 20, 21, 23	0	438	2.582
7	2, 4, 14, 18, 20, 21, 23	0	498	2.802
8	2, 4, 14, 18, 20, 21, 22, 25	1	567	2.982
9	4, 8, 13, 14, 18, 20, 21, 22, 25	2	628	3.140
10	4, 8, 13, 14, 17, 18, 20. 21, 22, 23	5	699	3.265
11	4, 5, 8, 10, 13, 14, 17, 21-24	9	743	3.352
12	4, 13, 14, 15, 17-24	16	837	3.443
13	4, 7, 13, 14, 15, 17-24	22	895	3.519
14	4, 6, 7, 13, 14, 15, 17-24	28	940	3.589
15	4, 6, 7, 13, 14, 15, 17-25	39	1013	3.639
16	2, 4, 6, 7, 13, 14, 15, 17-25	51	1084	3.684
17	1, 2, 4, 6, 7, 13, 14, 15, 17-25	69	1149	3.711
18	1, 2, 4, 6-9, 13-15, 17, 18, 20-25	97	1218	3.700
19	1-4, 6-9, 13-15, 17, 18, 20-25	119	1265	3.707
20	1-9, 13-15, 17, 18, 20-25	157	1322	3.673
21	1-9, 13-18, 20-25	198	1372	3.654
22	1-9, 13-25	243	1426	3.654
23	1-9, 11, 13-25	288	1463	3.625
24	1-9, 11-25	359	1502	3.563

Figure 52 Results from Confus/Recog Model for Moore Data Set

Subset size	Enumeration Scheme Selected Subsets	Confusion/ Recognition		Transmitted information
2	1, 17	0	130	1.000
3	2, 4, 14	0	213	1.585
4	2, 4, 14, 25	0	286	2.000
5	2, 4, 14, 22, 24	0	363	2.321
6	2, 4, 14, 20, 22, 24	0	430	2.584
7	2, 4, 14, 18, 20, 21, 25	0	492	2.803
8	2, 4, 14, 18, 20, 21, 22, 25	1	567	2.982
9	4, 8, 13, 14, 18, 20, 21, 22, 25	2	628	3.140
10	4, 8, 13, 14, 17, 18, 20, 21, 22, 23	5	699	3.265
11	4, 8, 13, 14, 17, 18, 20-24	10	774	3.364
12	1, 4, 8, 11, 13, 17, 18, 20-24	19	805	3.445
13	1, 4, 7, 13, 15, 17-24	22	889	3.520
14	4, 6, 7, 13-15, 17-24	28	940	3.589
15	4, 6, 7, 13-15, 17-25	39	1013	3.640
16	2, 4, 6, 7, 13-15, 17-25	51	1084	3.684
17	1, 2, 4, 6, 7, 13-15, 17-25	69	1149	3.711
18	1-4, 6, 7, 13-15, 17-25	97	1196	3.717
19	1-4, 6-8, 13-15, 17-25	129	1256	3.712
20	1-4, 6-9, 13-15, 17-25	170	1319	3.699
21	1-9, 13-15, 17-25	208	1376	3.673
22	1-9, 13-25	250	1426	3.655
23	1-9, 12-25	327	1473	3.627
24	1-9, 11-25	378	1510	3.600

Figure 53 Results from Enum Scheme Model for Moore Data Set

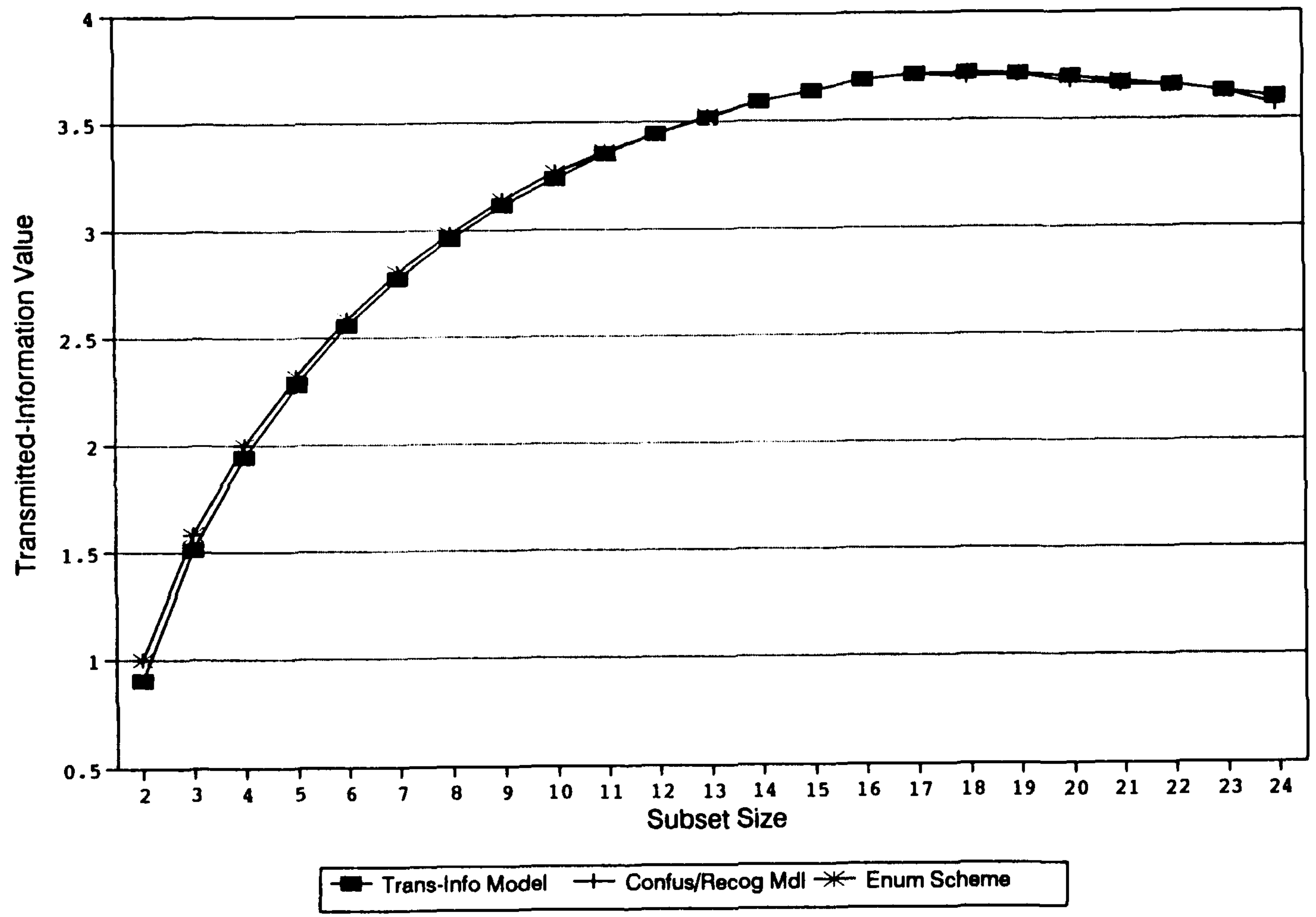

Figure 54 Moore Data Set: Transmitted-Information

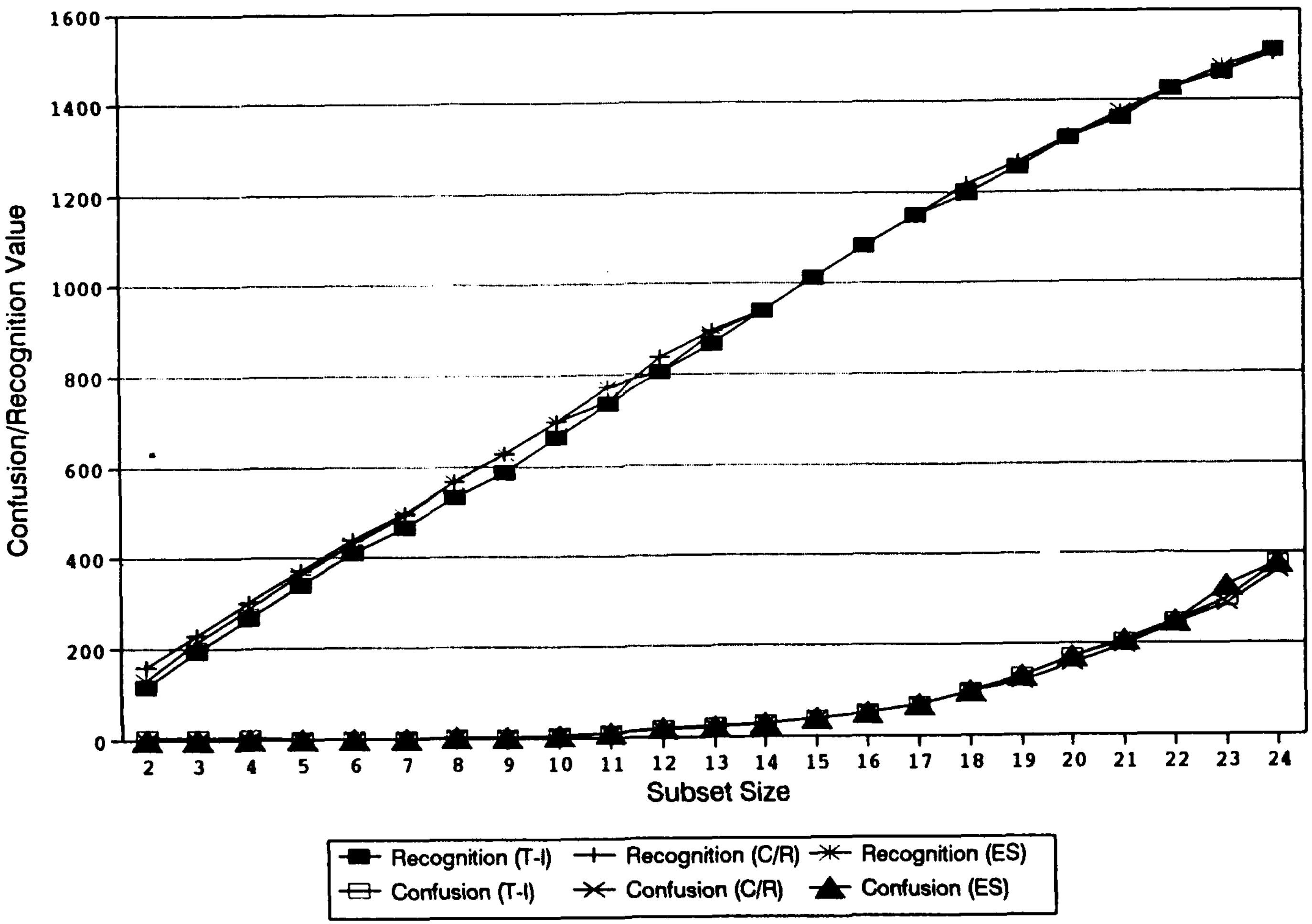

Figure 55 Moore Data Set: Confusion/Recognition

REFERENCES

Berg, S.A., *Introduction to Command, Control, and Communications (C3) Through Comparative Case Analysis*, Master's Thesis, Naval Postgraduate School, Monterey, California, March 1990.

Bethmann, R.C. and Malloy, K.A., *Command and Control: An Introduction*, Master's Thesis, Naval Postgraduate School, Monterey, California, March 1989.

Borland International, Inc., *Turbo Pascal Version 4.0 User's Guide*, 1987.

Bowen, H.M., Andreassi, J.L., Truax, S., and Orlansky, J., "Optimum Symbols for Radar Displays," *Human Factors*, v. 2, pp. 28-33, February 1960.

Brooke, A., Kendrick, D., and Meeraus, A., *GAMS: A User's Guide*, The Scientific Press, 1988.

Clarke, F.R., "Constant-Ratio Rule for Confusion Matrices in Speech Communication," *Journal of Acoustical Society of America*, v.29, pp. 715-720, 1957.

Fisher, D.L., "Optimal Performance Engineering: Good, Better, Best," *Human Factors*, in press.

Fourer, R., "Modeling Languages Versus Matrix Generators for Linear Programming," *ACM Transations on Mathematical Software*, v. 9, pp. 143-183, 1983.

Garner, W.R., *Uncertainty and Structure as Psychological Concepts*, John Wiley & Sons, Inc., 1962.

Gill, P.E., Murray, W., Murtagh, B.A., Saunders, M.A., and Wright, M.H., "GAMS/MINOS," *GAMS: A User's Guide*, The Scientific Press, 1988.

Green, P., and Pew, R.W., "Evaluating Pictographic Symbols: An Automotive Application," *Human Factors*, v. 20(1), pp. 103-114, February 1978.

Joint Chiefs of Staff, Command and Control Functional Analysis and Consolidation Review Panel, *Command and Control Functional Analysis and Consolidation Review Panel* Report, October 30, 1991.

Kantowitz, B.H. and Sorkin, R.D., *Human Factors: Understanding People-System Relationships*, John Wiley & Sons, Inc., 1983.

Marsten, R.E. and Singhal, J., "GAMS/ZOOM," *GAMS: A User's Guide*, The Scientific Press, 1988.

McQuitty, L.L., "Elementary Linkage Analysis for Isolating Orthoganal and Oblique Types and Typal Relevancies," *Educational and Psychological Measurements*, v. 17, pp. 207-229, 1957.

Moore, T.G., "Tactile and Kinaesthetic Aspects of Push-buttons," *Applied Ergonomics*, v. 5, pp. 66-71, 1974.

Orr, G.E., *Combat Operations C^3I: Fundamentals an Interactions*, Air University Press, July 1983.

Parsons, H.M., *Man-Machine Systems Experiments*, The Johns-Hopkins Press, 1972.

Pollack, I. and Decker, L., "Constant Confusions and the Constant Ratio Rule," *Language and Speech*, v. 3, no. 1, pp. 1-6, 1960.

Ravindran, A., Phillips, D.T., and Solberg, J.J., *Operations Research: Principles and Practice*, 2nd ed., 1987.

Rohl, J.S., *Recursion Via PASCAL*, Cambridge University Press, 1983.

Romesburg, H.C., *Cluster Analysis for Researchers*, Lifetime Learning Publications, 1984.

Sanders, M.S. and McCormick, E.J., *Human Factors in Engineering and Design*, 6th ed., McGraw-Hill Book Co., 1987.

Sunset Software Technology, *XA Professional Linear Programming System*, 1987.

Theise, E.S., "Finding a Subset of Stimulus-Response Pairs with Minimum Total Confusion: A Binary Integer Programming Approach," *Human Factors*, v. 31(3), pp.291-305, June 1989.

Wickens, C.D., *Engineering Psychology and Human Performance*, Merrill, 1984.

Wilpon, J.G., "A Study on the Ability to Automatically Recognize Telephone-Quality Speech From Large Customer Populations," *AT&T Technical Journal*, v. 64, no. 2, pp. 423-451, February 1985.

INITIAL DISTRIBUTION LIST

1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, California 93943-5000	2
3.	Director for Command, Control, and Communications Systems, Joint Staff Washington, DC 20318-6000	1
4.	C3 Academic Group, Code CC Naval Postgraduate School Monterey, California 93943-5000	1
5.	AFIT/NR Wright-Patterson AFB, Ohio 45433-6583	1
6.	AFIT/CIRK Wright-Patterson AFB, Ohio 45433-6583	1
7.	Professor Eric S. Theise Operations Research Department, Code ORTH Naval Postgraduate School Monterey, California 93943-5000	3
8.	Capt Michael J. Sheehan TIC/APT Scott AFB, IL 62225	1